

Advance in Fireworks Algorithm and its Applications

Ying Tan (谭营)
Peking University

Contact

ytan@pku.edu.cn
www.cil.pku.edu.cn



This PPT is available at www.cil.pku.edu.cn/research/fa.

Outlines

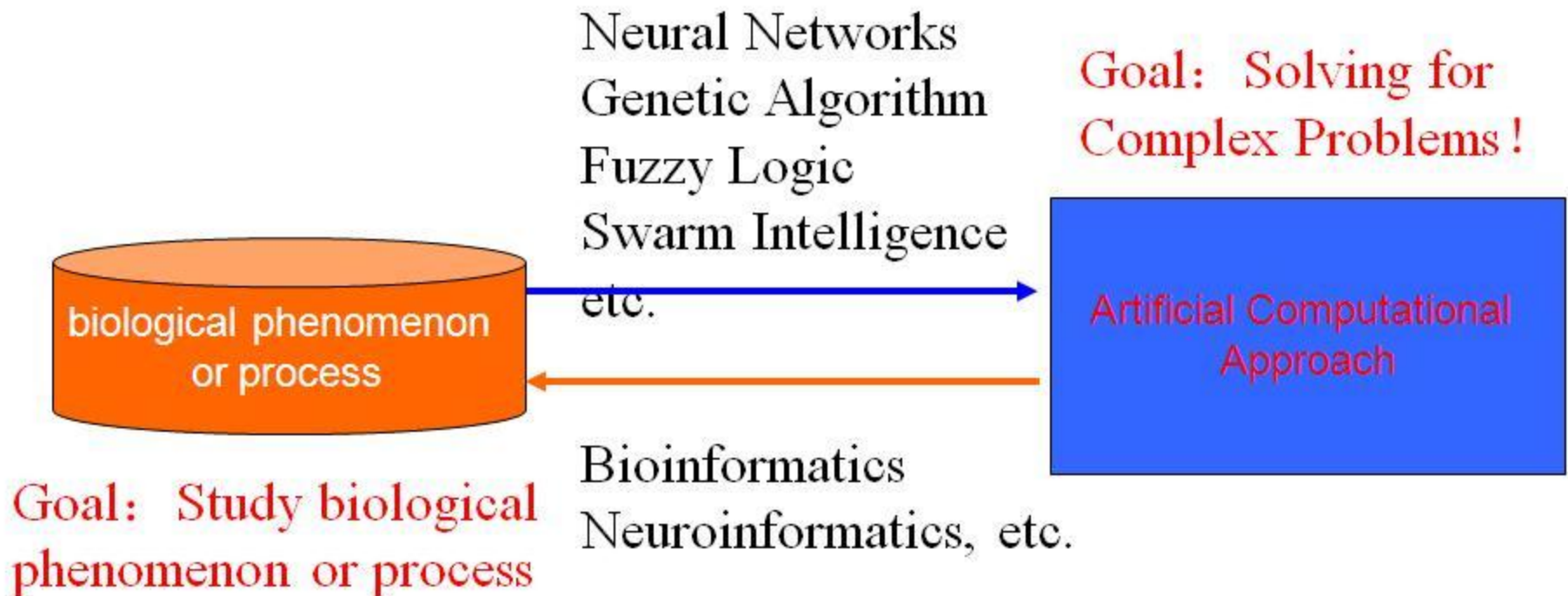
1. Brief Introduction to Swarm Intelligence
2. basic Fireworks Algorithm (FA)
3. FA Variant
4. Latest Applications of FA
 - FA for NMF
 - FA for Document Clustering
5. Concluding Remarks

1. Brief Introduction to Swarm Intelligence

1.1 **Swarm Intelligence (SI)** refers to

- Simple individuals and their information processing process
- Interaction between individuals or with environment

*from nature to
artificial systems*



1. Brief Introduction to SI

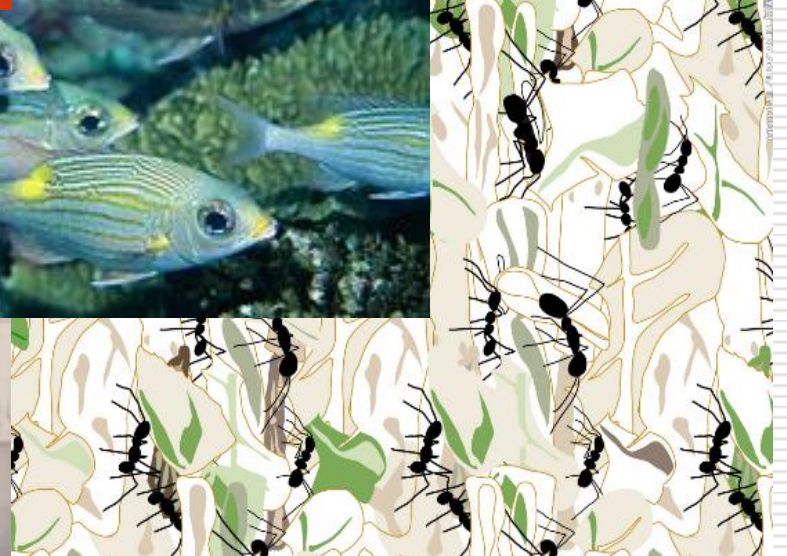
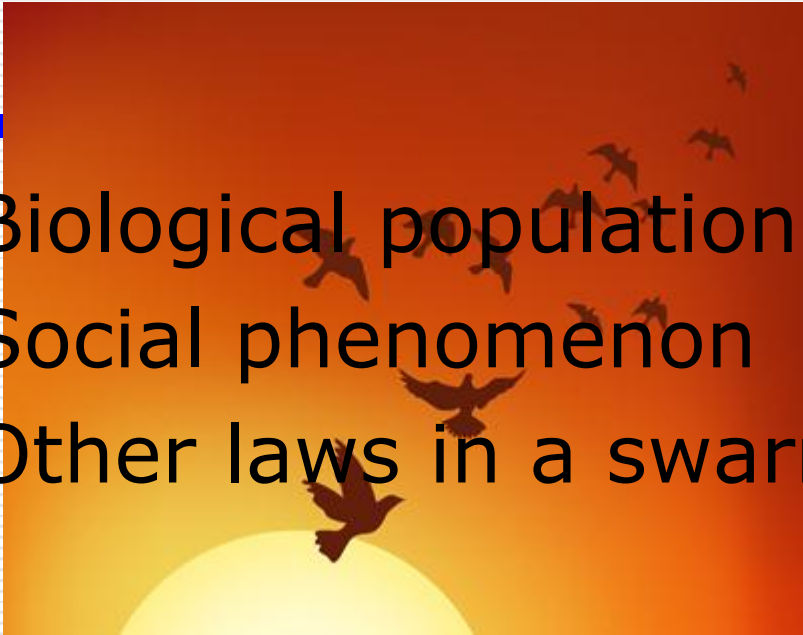
1.2 Some Famous SI Algorithms:

- Particle Swarm Optimization (PSO)
- Ant Colony Optimization (ACO)
- Artificial Immune System (AIS)
- Bee Colony Optimization (BCO)
- Bacterial Foraging Optimization (BFO)
- Fish School Search (FSS)
- Seeker Optimization Algorithm (SOA)

To name a few

1.

- ☐ Biological population
- ☐ Social phenomenon
- ☐ Other laws in a swarm in nature



1.2.1 Particle Swarm Optimization (PSO)

Inspired by the
search food
of flocks



1.2.1 Particle Swarm Optimization

- A birds flock is searching for a food, and every bird does not know where the food is. But, they know presently the distance of each bird to the food

**how to make a strategy
that the bird can get to the
food fastest?**

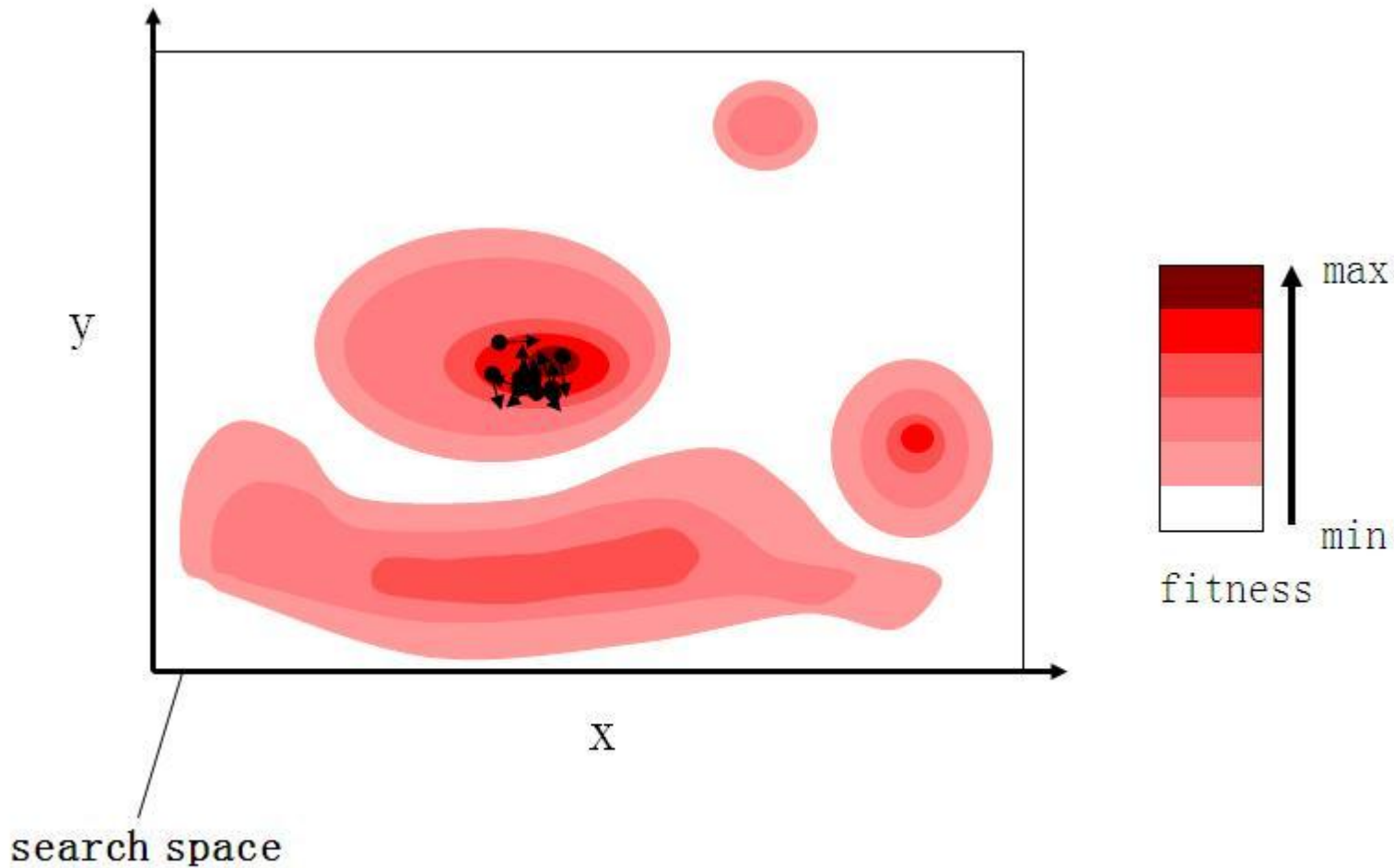
This seeking behavior was associated with that of an optimization

1.2.1 PSO principle



How to choose ?

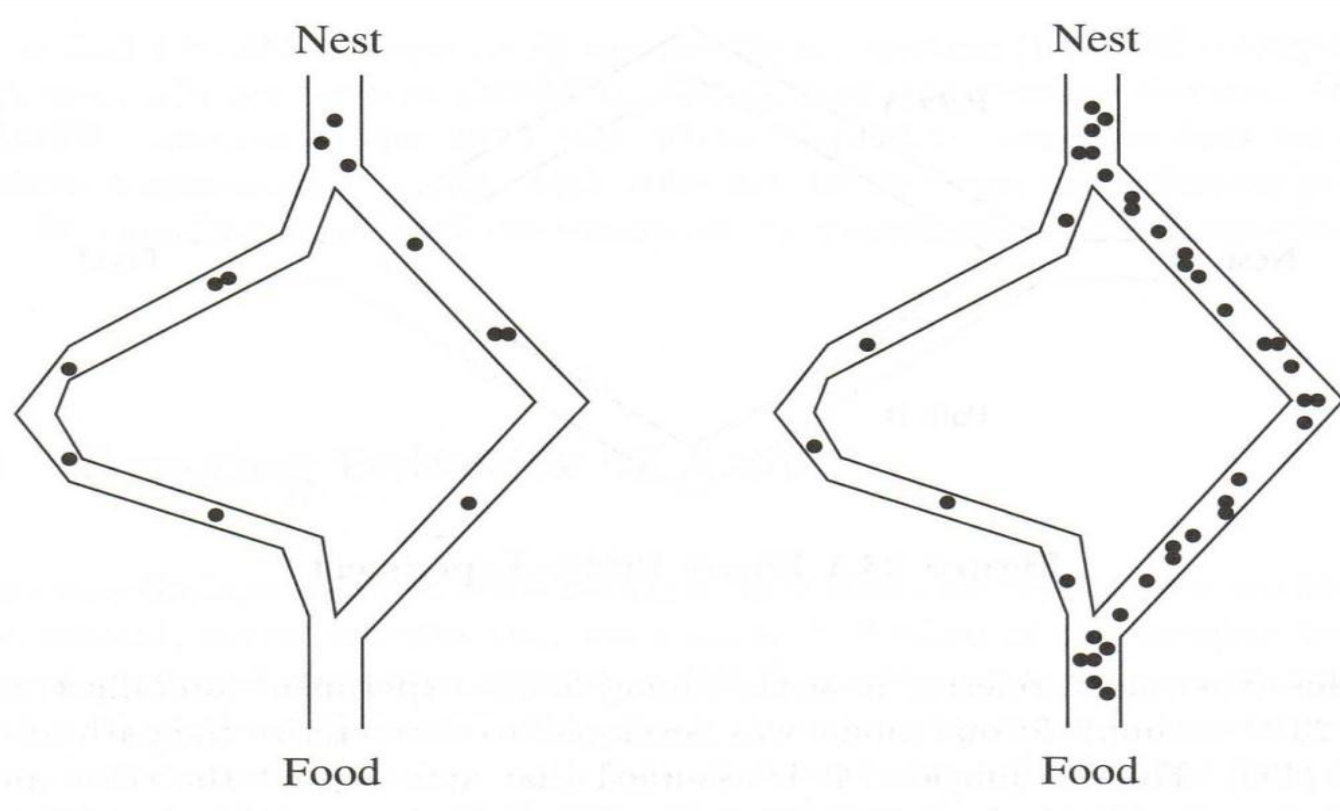
1.2.1 Visual demonstration of PSO



1.2.2 Ant Colony Optimization (ACO)

- Ant system searches Food from Nest
- ANT.EXE (AVI1)

Double Bridge *Experiment*



Auto-catalytic (positive feedback) process



Sponsors and Co-sponsors: Shenzhen University, Peking University, Xi'an Jiaotong-Liverpool University, and Natural Science Foundation of China.
Technical Co-sponsors: IEEE Computational Intelligence Society, International Neural Network Society, World Federation on Soft Computing, Springer, etc.

CALL FOR PAPERS

The third International Conference on Swarm Intelligence (ICSI'2012) will be held in Shenzhen, China from June 17 to 20, 2012. Shenzhen is a coastal city located in the south of China's Guangdong Province, adjoining Hong Kong. As China's first Special Economic Zone, Shenzhen has been a touchstone for the country's reform and opening-up policy. In Shenzhen, you will behold the epitome of China's achievements in reform and opening during the past three decades. Shenzhen also has built a relationship with its neighbor Hong Kong based on promoting cross-border tour activities. Frequent and convenient transportation between these two cities will offer you a fabulous opportunity to experience two cities in just a single day.

ICSI'2012 serves as a forum for scientists, engineers, educators and practitioners to exchange the latest advantages in theory, technologies, and applications of swarm intelligence. Prospective authors are invited to contribute high-quality papers (6-10 pages) to ICSI'2012 through [Online Submission System](#). Papers presented at ICSI'2011 will be published in [Springer Lecture Notes in Computer Science](#) (indexed by EI, ISTP, DBLP, ISI) and some high-quality papers will be selected for special issues in SCI-indexed International Journals. Topics of interest include, but are not limited to:

Theories	Algorithms	Models	Applications
<ul style="list-style-type: none"> •Swarm-based optimization techniques •Swarm computing •Particle swarm optimization •Ant colony optimization •Fish school search •Stochastic diffusion search (SDS) •Swarm robotics •Artificial life •Bioinformatics •Cognitive science •Social evolution •Optimization theory and methods •Grammatical Swarm •Evolutionary computation •Grammatical Evolution •Natural computing •Simulation and emulation of nature •Multi-agent based complex systems •Collective intelligence •Social intelligence •Social computing 	<ul style="list-style-type: none"> •PSO algorithms •ACO algorithms •FSS algorithms •Bees algorithm •Artificial Bee Colony Algorithms •Cultural algorithms •Genetic algorithms •Co-evolution algorithms •Memetic algorithms •Hybrid optimization methods •Evolutionary programming •Evolutionary strategy •Hybridization method •Evolutionary learning systems <p>Emerging areas</p> <ul style="list-style-type: none"> •Molecular Computing •DNA computing •Quantum computing •Granularity computing 	<ul style="list-style-type: none"> •Computational swarm models •Artificial immune system •Multi-objective optimization •Evolutionary intelligent agents •Constrained environment •Information System •Analytic models of emergent behaviors •Classifier system •Evolving fuzzy system •Evolving neural networks •Models for social insets •Bio-inspired computing models •Socio-inspired computing models •Hybridizations of algorithms 	<ul style="list-style-type: none"> •Pattern recognition •Intelligent control •Financial prediction •Information security •Web intelligence •Data mining •2D or 3D virtual swarms •Virtual reality •Audio or Image processing •Telecommunications systems •Evolutionary robotics •Automatic control •Routing •Socio-technical systems •Scheduling •Cybernetics and self-organization •Intelligent traffics •Simulations •Electronics systems •Defense security •Video surveillance

Important Dates:

Special session proposals deadline:

October 01, 2011

Paper submission deadline:

December 01, 2011

Notification of acceptance:

March 01, 2012

Camera-ready copy and author registration:

March 31, 2012

Welcome to ICSI 2012 Shenzhen, CN.

www.ic-si.org

General Chairs

Ying Tan, PKU, China
 Ji Zhen, SZU, China

Program Committee Chairs

Yuhui Shi, XUTLU, China

Advisory Committee Chairs

Guoliang Chen, SZU, China
 Russell C. Eberhart, Indiana, USA

Member of Advisory Committee

Hojjat Adeli, OSU, USA
 Walter J. Freeman, UCB, USA
 Xingui He, PKU, China
 Lei Guo, CAS, China
 Jong-Hwan Kim, KAIST, S.Korea
 Gary G. Yen, Okstate, USA
 Deyi Li, ICESE, China
 Yanda Li, Tsinghua Univ, China
 Zushu Li, CQU, China
 Leszek Rutkowski, CUT, Poland
 Benjamin W. Wah, CUHK, HK
 Qidi Wu, Tongji Univ, China
 Yixin Zhong, BUPT, China

Technical Committee Chairs

Zexuan Zhu, SZU, China
 Martin Middendorf, ULeipzig, Germany
 Qinghua Wu, ULiverpool, UK
 Yew-Soon Ong, NTU, Singapore

Special Session Chairs

Shan He, UBirmingham, UK

Publication Chairs

Rajkumar Roy, Ucranfield, UK
 Radu-Emil Precup, UPT, Romania

Publicity Chairs

Xiaodong Li, RMIT Univ, Australia
 Haibo He, URI, USA
 Zhuohong You, SZU, China

Finance and Registration Chairs

Andreas Janacek, UVienna, Austria
 Chao Deng, PKU, China

Local Arrangement Chairs

Jiarui Zhou, SZU, China



1.3 Introduction to Fireworks Algorithm

When a firework is set Off, a shower of sparks will fill the local space around the firework.

□ *The explosion can be viewed as a search in the local Space around a firework.*

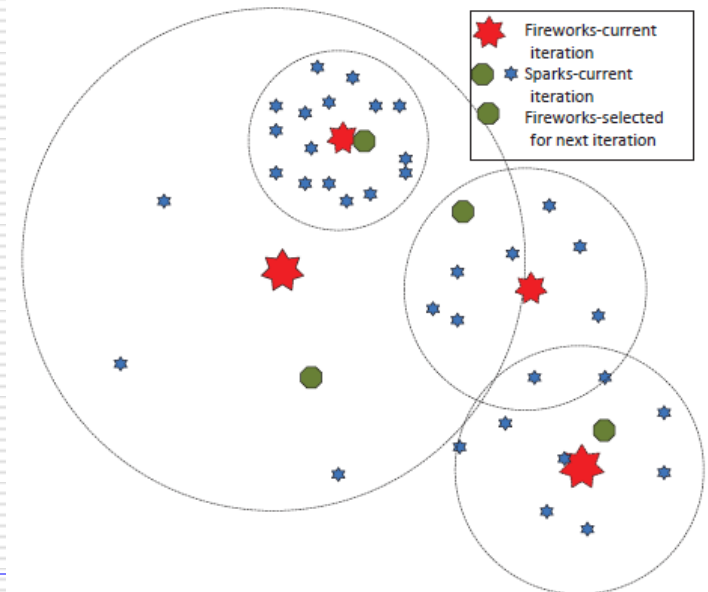
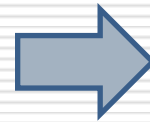


Fig. 1. Search Process of Fireworks Algorithm



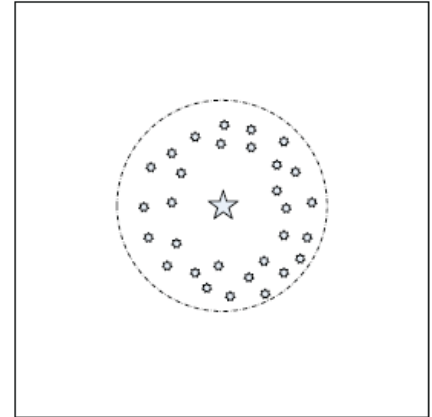
from Nature to Mathematical Modeling

□ Definition of firework

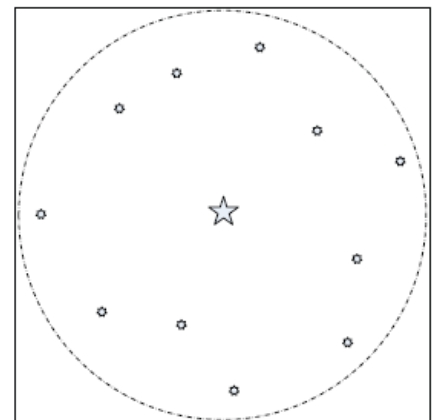
Good firework: firework can generate a big population of sparks within a small range.

Bad firework: firework that generate a small Population of sparks within a big range.

The basic FA algorithm is on the basis of simulating the process of the firework explosion illuminating the night sky.



(a) Good explosion



(b) Bad explosion

Fig. 2. Two types of fireworks explosion

2. Basic Firework Algorithm (FA)

2.1 Problem description

2.2 Flowchart of FA

2.3 Design of basic FA

- Selection operators

- Number of son sparks

2.4 Experimental results

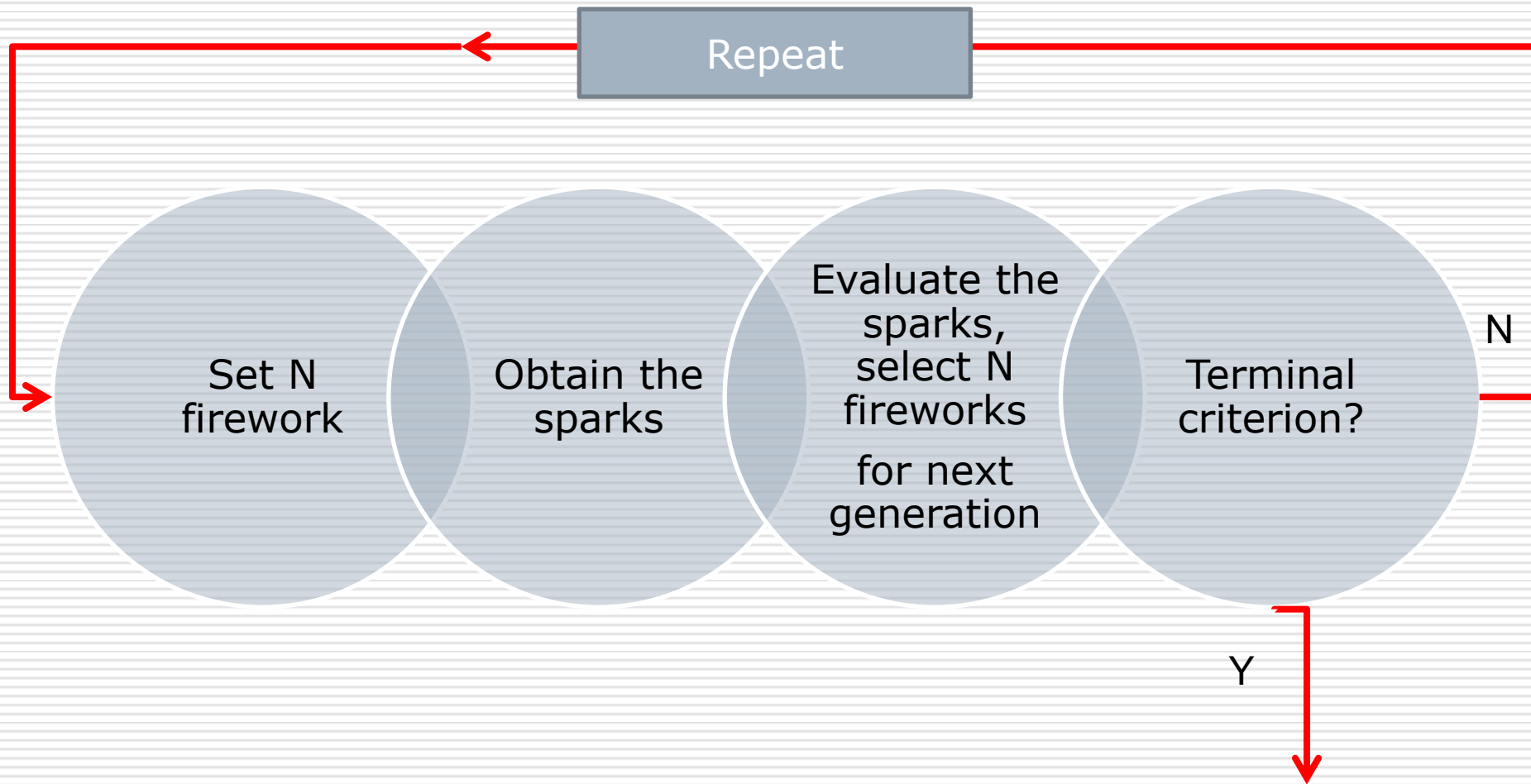
2.1 Problem description

- Suppose the FA is designed for the general optimization problems:

$$\text{Minimize } f(x) \in \mathbb{R}, x_{min} \leq x \leq x_{max},$$

Where $x = x_1, x_2, \dots, x_d$ denotes a location in the potential space, $f(x)$ is an objective function, and x_{min} and x_{max} denote the bounds of the potential space.

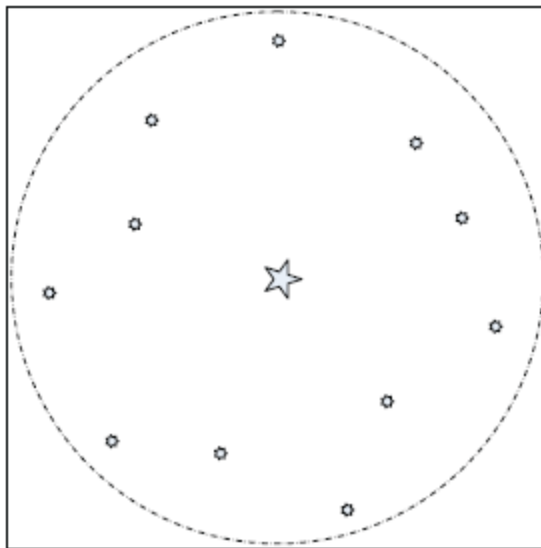
2.2 FA's Flowchart



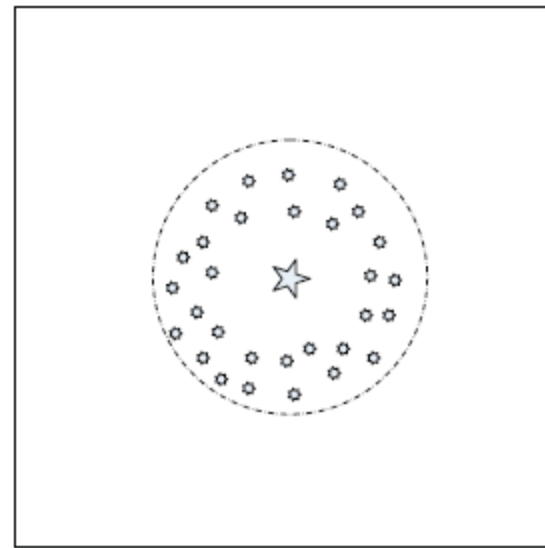
2.3 Design of basic Firework Algorithm

Number of sparks generated by each firework X_i is defined as follows:

$$s_i = m * \frac{y_{max} - f(x_i) + \zeta}{\sum_{i=1}^n (y_{max} - f(x_i)) + \zeta}$$



bad firework



good firework

2.3 Design of basic Firework Algorithm

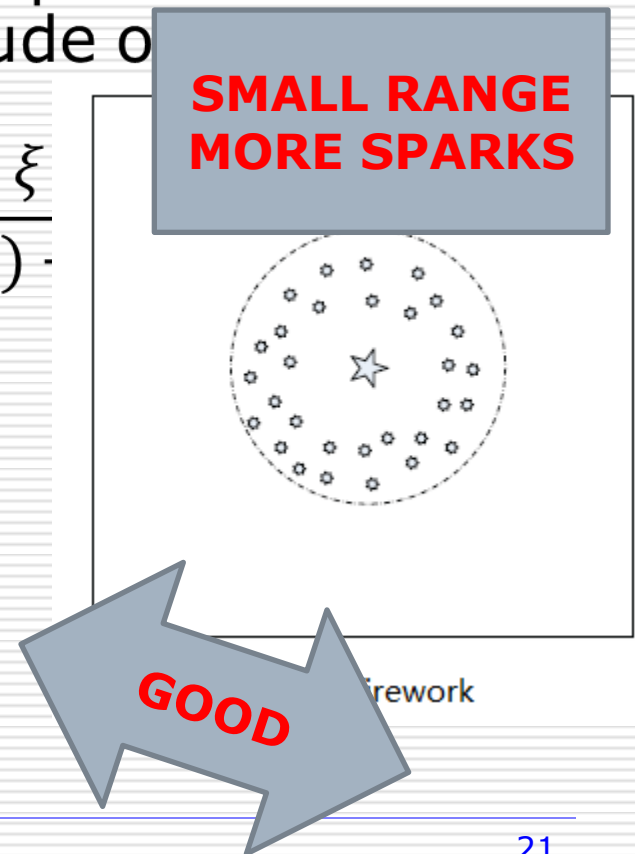
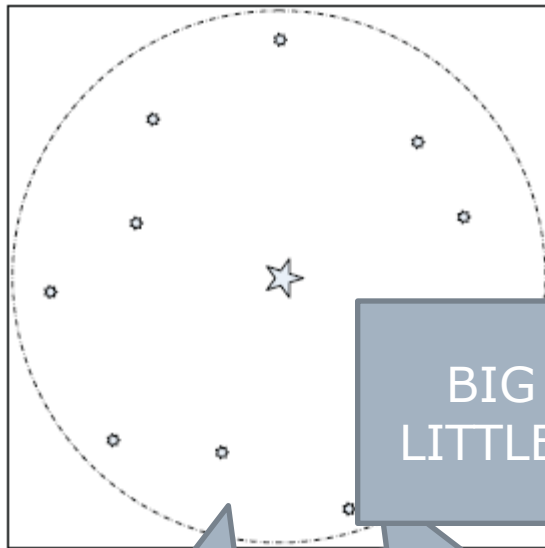
To avoid overwhelming effects of splendid fireworks, bounds are defined

$$\hat{s}_i = \begin{cases} \text{round}(a * m) & \text{if } s_i < am \\ \text{round}(b * m) & \text{if } s_i < am, a < b < 1 \\ \text{round}(s_i) & \text{otherwise} \end{cases}$$

2.3 Design of basic Firework Algorithm

In contrast to the design of sparks number, the amplitude of a good firework explosion is smaller than that of a bad one. Amplitude is defined by:

$$A_i = \frac{f(x_i) - y_{min} + \xi}{\sum_{i=1}^n (f(x_i) - y_{min})}$$



2.3 Design of basic Firework Algorithm

Generating Sparks. In explosion, sparks may undergo the effects of explosion from random z directions (dimensions).

Algorithm 1. Obtain the location of a spark

Initialize the location of the spark: $\tilde{x}_j = x_i$;

$z = \text{round}(d \cdot \text{rand}(0, 1))$;

Randomly select z dimensions of \tilde{x}_j ;

Calculate the displacement: $h = A_i \cdot \text{rand}(-1, 1)$;

for each dimension $\tilde{x}_k^j \in \{\text{pre-selected } z \text{ dimensions of } \tilde{x}_j\}$ **do**

$\tilde{x}_k^j = \tilde{x}_k^j + h$;

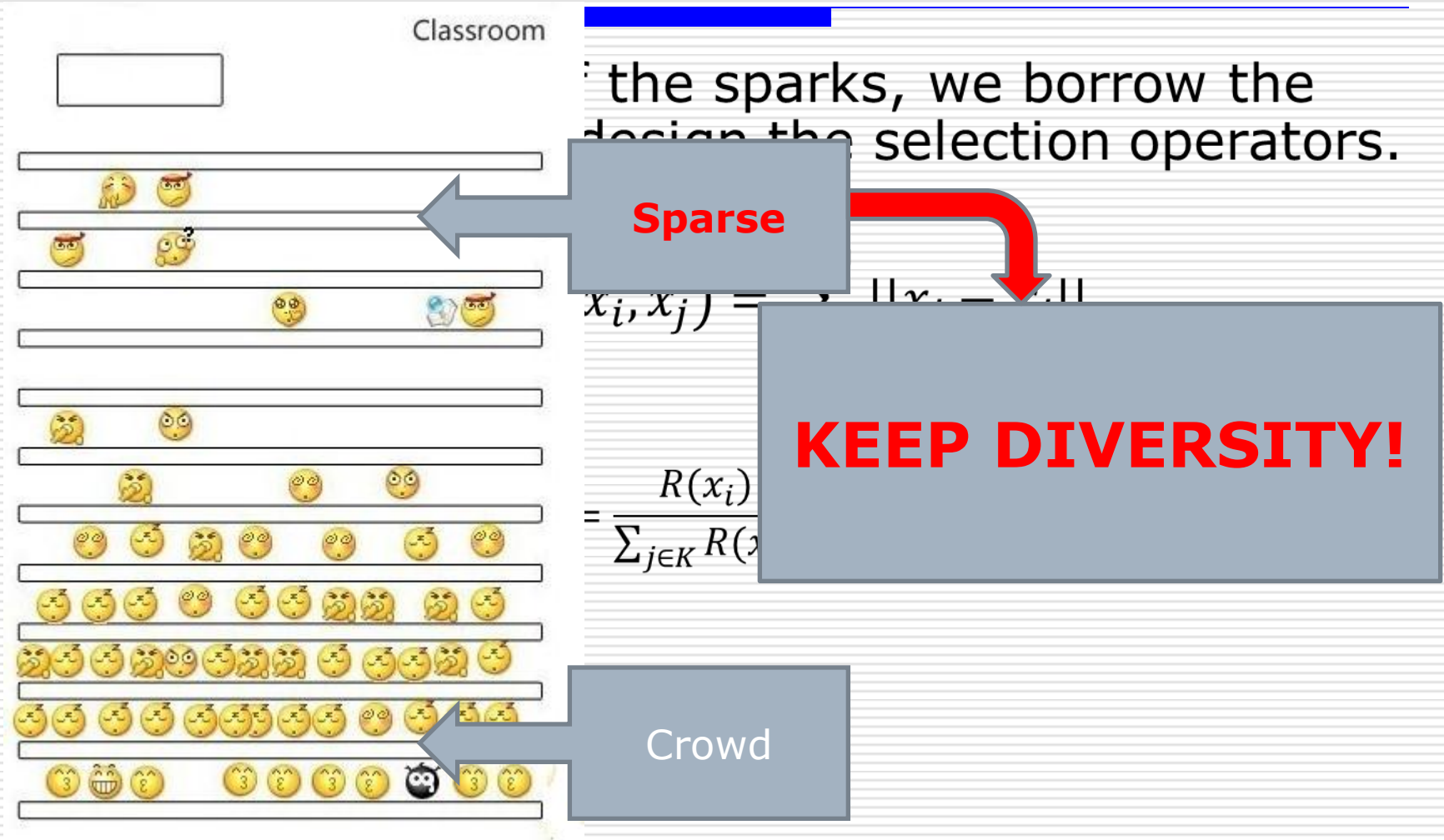
if $\tilde{x}_k^j < x_k^{\min}$ or $\tilde{x}_k^j > x_k^{\max}$ **then**

 map \tilde{x}_k^j to the potential space: $\tilde{x}_k^j = x_k^{\min} + |\tilde{x}_k^j| \% (x_k^{\max} - x_k^{\min})$;

end if

end for

2.3 Design of basic Firework Algorithm



2.3 Design of basic Firework Algorithm

To keep the diversity of sparks, we design another way of generating sparks----Gaussian explosion.

Algorithm 2. Obtain the location of a s

Initialize the location of the spark: $\hat{x}_j = x$
 $z = \text{round}(d \cdot \text{rand}(0, 1))$;

Randomly select z dimensions of \hat{x}_j ;

Calculate the coefficient of Gaussian explo

for each dimension $\hat{x}_k^j \in \{\text{pre-selected } z$

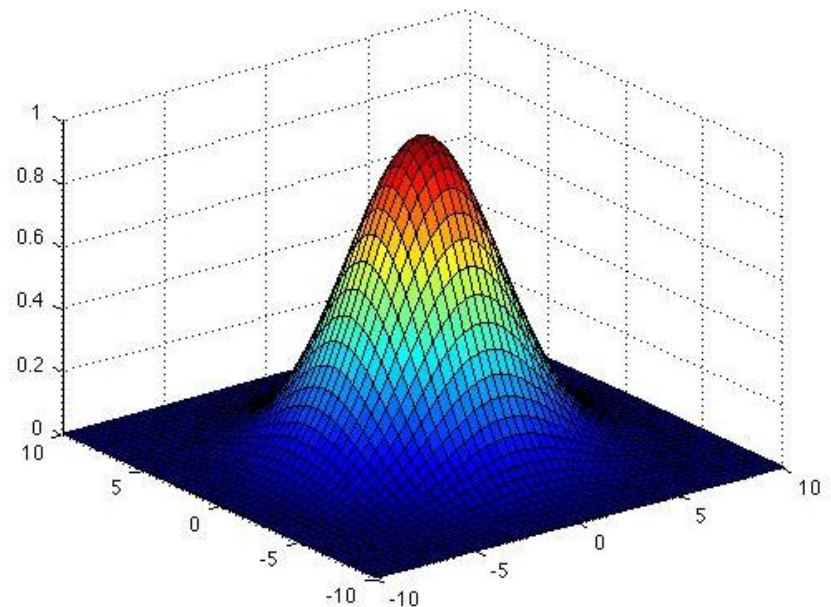
$\hat{x}_k^j = \hat{x}_k^j \cdot g$;

if $\hat{x}_k^j < x_k^{\min}$ or $\hat{x}_k^j > x_k^{\max}$ **then**

map \hat{x}_k^j to the potential space: $\hat{x}_k^j = :$

end if

end for

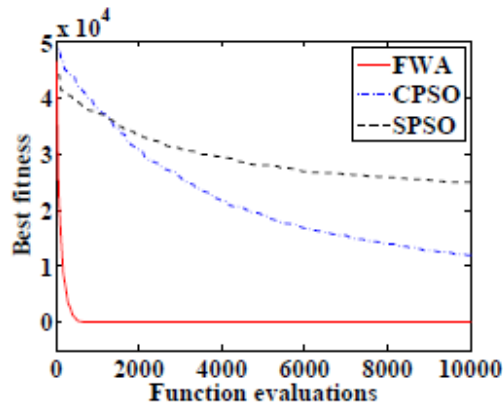


2.4 Experiments results of basic FA

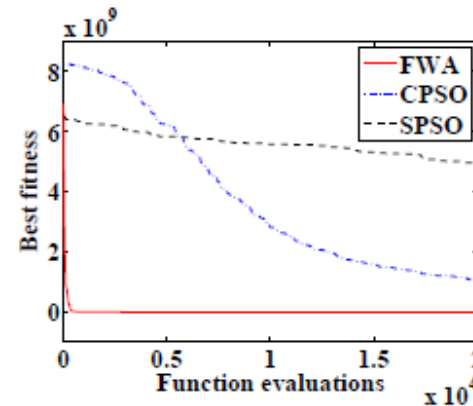
Table 1. Nine benchmark functions utilized in our experiments. The details, including the expression of the functions, feasible bounds, and initialization intervals, are given in the table.

Function	Expression	Feasible bounds	Initialization	D
Sphere	$F_1 = \sum_{i=1}^D x_i^2$	$[-100, 100]^D$	$[30, 50]^D$	30
Rosenbrock	$F_2 = \sum_{i=1}^{D-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	$[-100, 100]^D$	$[30, 50]^D$	30
Rastrigrin	$F_3 = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10)$	$[-100, 100]^D$	$[30, 50]^D$	30
Griewank	$F_4 = 1 + \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos(\frac{x_i}{\sqrt{i}})$	$[-100, 100]^D$	$[30, 50]^D$	30
Ellipse	$F_5 = \sum_{i=1}^D 10^4 \frac{i-1}{D-1} x_i^2$	$[-100, 100]^D$	$[15, 30]^D$	30
Cigar	$F_6 = x_1^2 + \sum_{i=2}^D 10^4 x_i^2$	$[-100, 100]^D$	$[15, 30]^D$	30
Tablet	$F_7 = 10^4 x_1^2 + \sum_{i=2}^D x_i^2$	$[-100, 100]^D$	$[15, 30]^D$	30
Schwefel	$F_8 = \sum_{i=1}^D ((x_1 - x_i^2)^2 + (x_i - 1)^2)$	$[-100, 100]^D$	$[15, 30]^D$	30
Ackley	$F_9 = 20 + e - 20 \exp \left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} \right) - \exp \left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i^2) \right)$	$[-100, 100]^D$	$[15, 30]^D$	30

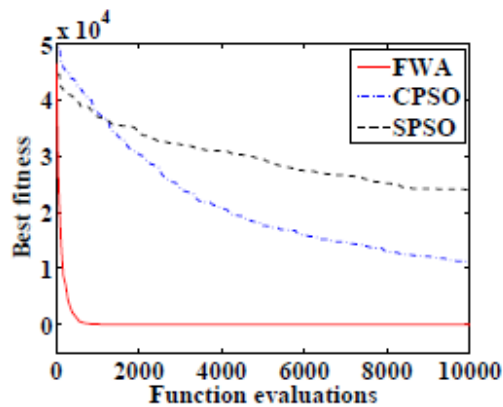
2.4 Experiments results of basic FA



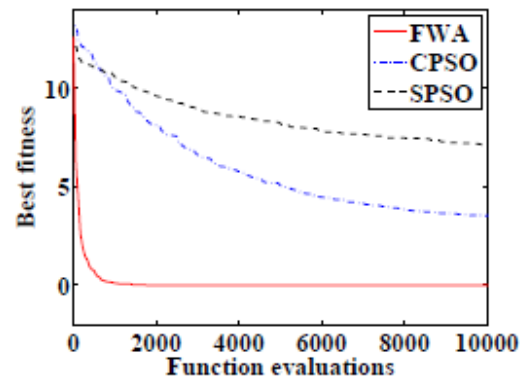
(a) Sphere



(b) Rosenbrock

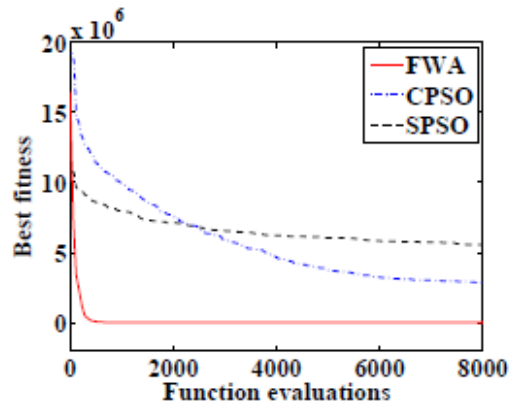


(c) Rastrigin

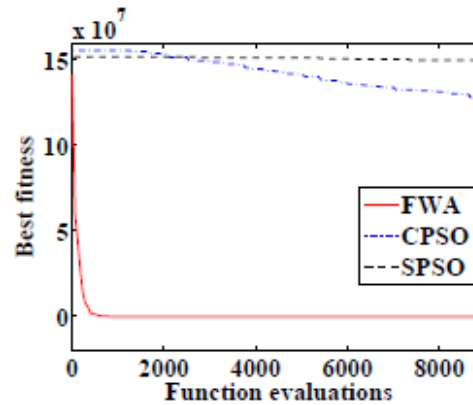


(d) Griewank

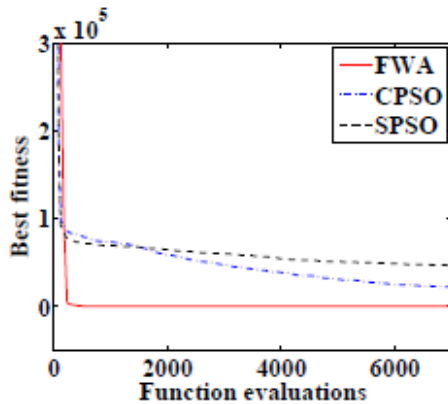
2.4 Experiments results of basic FA



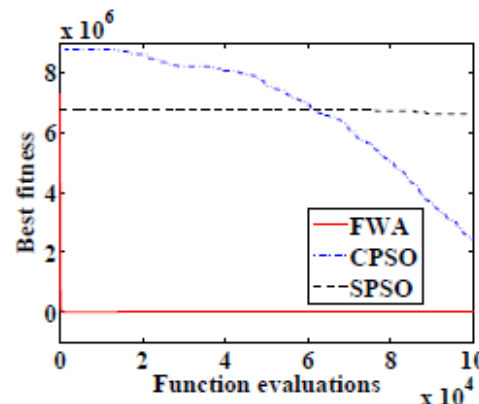
(e) Ellipse



(f) Cigar



(g) Tablet



(h) Schwefel

2.4 Experiments results of basic FA

Table 2. Statistical mean and standard deviation of solutions found by the FA, the CPSO and the SPSO on nine benchmark functions over 20 independent runs

Function	Function evaluations	FA's mean \pm StD	CPSO's mean \pm StD	SPSO's mean \pm StD
Sphere	500000	0.000000 \pm 0.000000	0.000000 \pm 0.000000	1.909960 \pm 2.594634
Rosenbrock	600000	9.569493 \pm 12.12829	33.403191 \pm 42.513450	410.522552 \pm 529.389139
Rastrigrin	500000	0.000000 \pm 0.000000	0.053042 \pm 0.370687	167.256119 \pm 42.912873
Griewank	200000	0.000000 \pm 0.000000	0.632403 \pm 0.327648	2.177754 \pm 0.294225
Ellipse	500000	0.000000 \pm 0.000000	0.000000 \pm 0.000000	53.718807 \pm 68.480173
Cigar	600000	0.000000 \pm 0.000000	0.000000 \pm 0.000000	0.002492 \pm 0.005194
Tablet	500000	0.000000 \pm 0.000000	0.000000 \pm 0.000000	1.462832 \pm 1.157021
Schwefel	600000	0.000000 \pm 0.000000	0.095099 \pm 0.376619	0.335996 \pm 0.775270
Ackley	200000	0.000000 \pm 0.000000	1.683649 \pm 1.317866	12.365417 \pm 1.265322

Explosion Search Demonstrations of FA

- Sphere (AVI2)
- Benchmark function f18 (AVI3)

3. FA variants

3.1 Motivation of FA variants

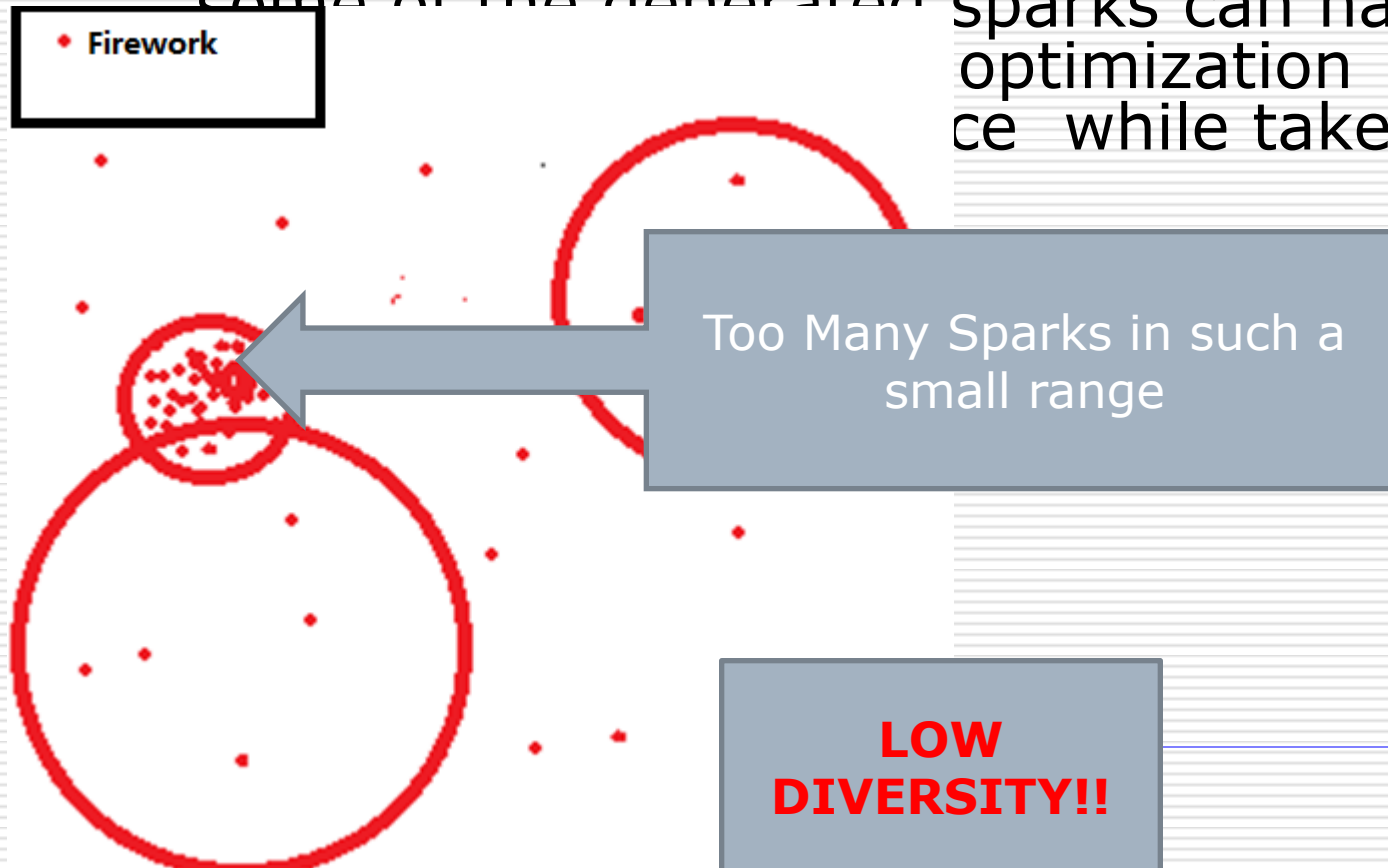
3.2 Definition for FA variants

3.3 FA with fixed minimal explosion
amplitude

3.4 FA with dynamic minimal explosion
amplitude

3.1 Motivation of FA variants

- By comprehensively analyzing the techniques of fireworks explosion, our findings suggest that some of the generated sparks can hardly make optimization if they do ce while take a lot



3.2 Definition for FA variants

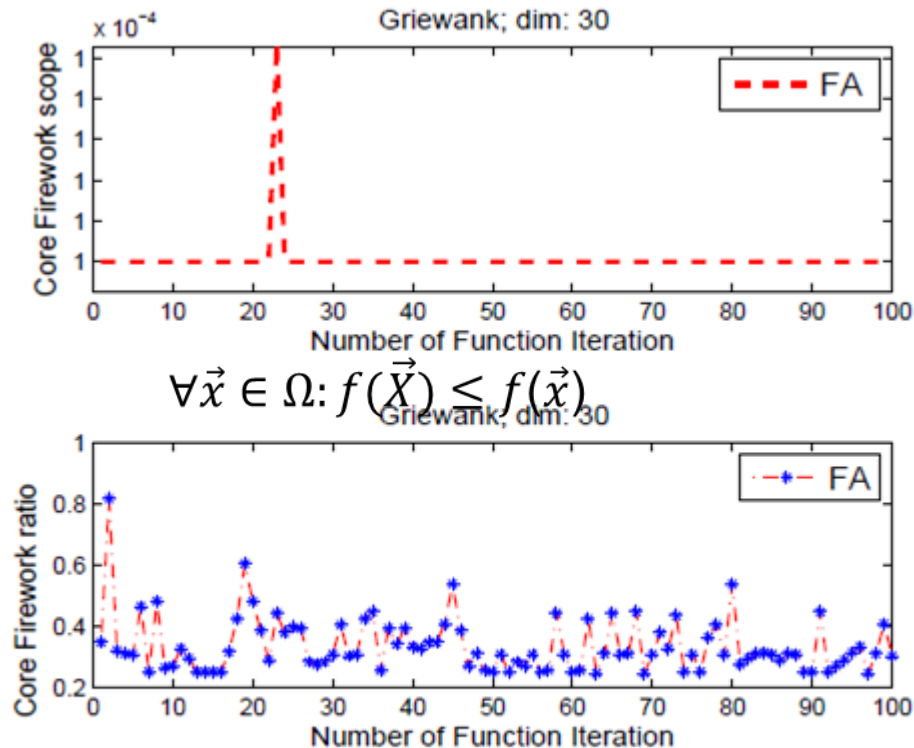


Fig. 4. Analysis of the Core Firework Explosion Amplitude and Ratio

Core Firework: In iterative process, there is fixed number of fireworks which can generate sparks. Among the fireworks in each iteration, the firework with minimal fitness is defined as Core Firework.

3.2 Definition for FA variants

Significance Improvement

whether the Core Firework
we define that if it generates
generation whose fitness
iteration as Significance I

$$Signif(x) = \begin{cases} 1 & , \left\{ \left| \frac{f_{i+1} - f_i}{f_{i+1}} \right| > x \cap f_{i+1} \leftarrow Core\ Firework \right\} \\ 0 & , others \end{cases}$$

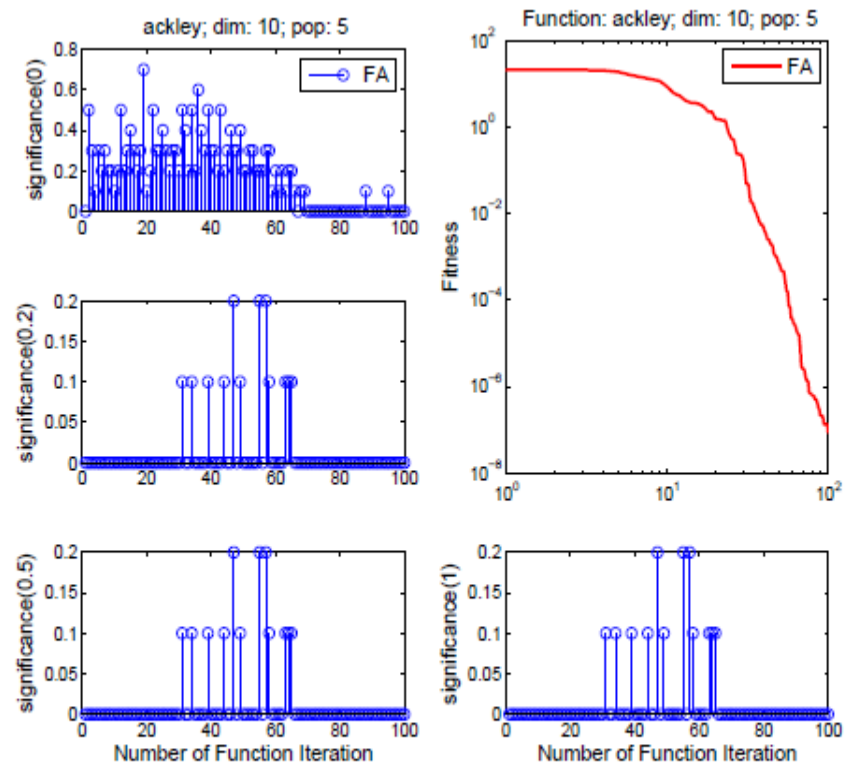


Fig. 5. Analysis of the Core Firework Explosion Significance Improvement

3.2 Definition for FA variants

Local Minimal Space: Given a function $f: \Omega \subseteq R^n \rightarrow R$, in a continual space $\Psi \subseteq \Omega$, there $\exists x, \exists \varepsilon$, and $f(x) - f(X) \geq 0$, when $|x - X| \leq \varepsilon$, then Ψ is a local minimal space.

So optimization problems can be classified as :only one local local minimal space problems (**Uni-modal**), several local minimal space problems (**Multi-modal**) and many local minimal space problems (**Hybrid-modal**).

3.3 FA with fixed minimal explosion

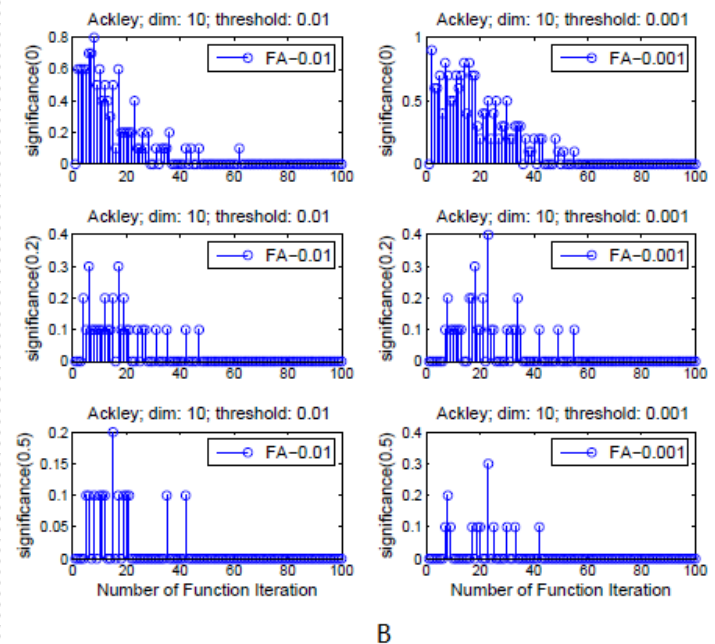
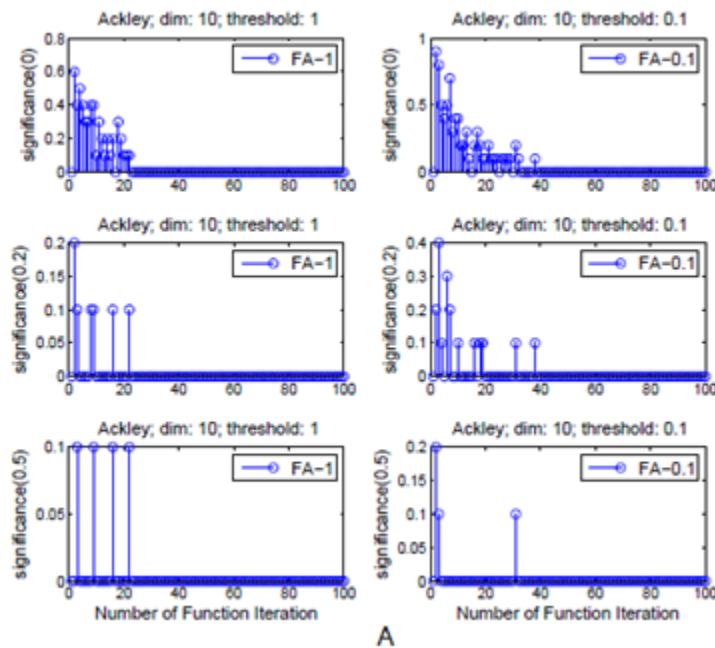
amplitude

```

1: Initialization.
2: Randomly select  $n$  fireworks at  $n$  locations
3: for  $t = 1$  to  $Iter$  do
4:   evaluate the  $n$  fireworks using object function
5:   calculate the number of sparks of each firework
6:   calculate the explosion amplitude of each firework
7:   for  $i = 1$  to  $N$  do
8:     compute current fitness  $f(X_i)$ 
9:     if explosion amplitude of Core Firework  $< \mathbf{Thresh-}$ 
       old then
10:       set it as Threshold
11:     end if
12:     generate the sparks for each firework
13:     obtain the Gaussian Mutation of all sparks
14:     obtain the optimal spark among the SET of all
       sparks, the SET include generated sparks, Gaussian
       Mutation sparks and  $n$  fireworks
15:     randomly select other  $n - 1$  fireworks
16:     obtain  $n$  fireworks for next iteration
17:   end for
18: end for
19: return the optimal solutions found by the swarm
  
```

3.3 FA with fixed minimal explosion amplitude

Fig. 6. Analysis of the Core Firework Explosion Significance Improvement with fixed minimal explosion amplitude



3.3 FA with fixed minimal explosion amplitude

- ❑ As can be seen, the bigger fixed minimal scope can accelerate the process of searching the optimal solution at the first iterations. However, when it comes to the local minimal space, the higher the fixed minimal scope, the harder it searches the optimal solution.
- ❑ It would seem that function Ackley is uni-modal function. And multi-modal function is the same, the difference is that only when the swarm comes to a local minimal space rather than a global minimal space compared to uni-modal functions.

3.4 FA with dynamic minimal explosion amplitude

- Experiments with fixed dynamic explosion amplitude suggest that FA should be loaded with a higher explosion amplitude while it does Not come to a local minimal space, thus to enhance the possibility of obtain a better solution. When it got to a local minimal space, then the fireworks begin to search the Space within a small range from the found good solution.

3.4 FA with dynamic minimal explosion amplitude

**If FA have not get
any better
positions in N
consecutive steps?**



57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

11/1/2011

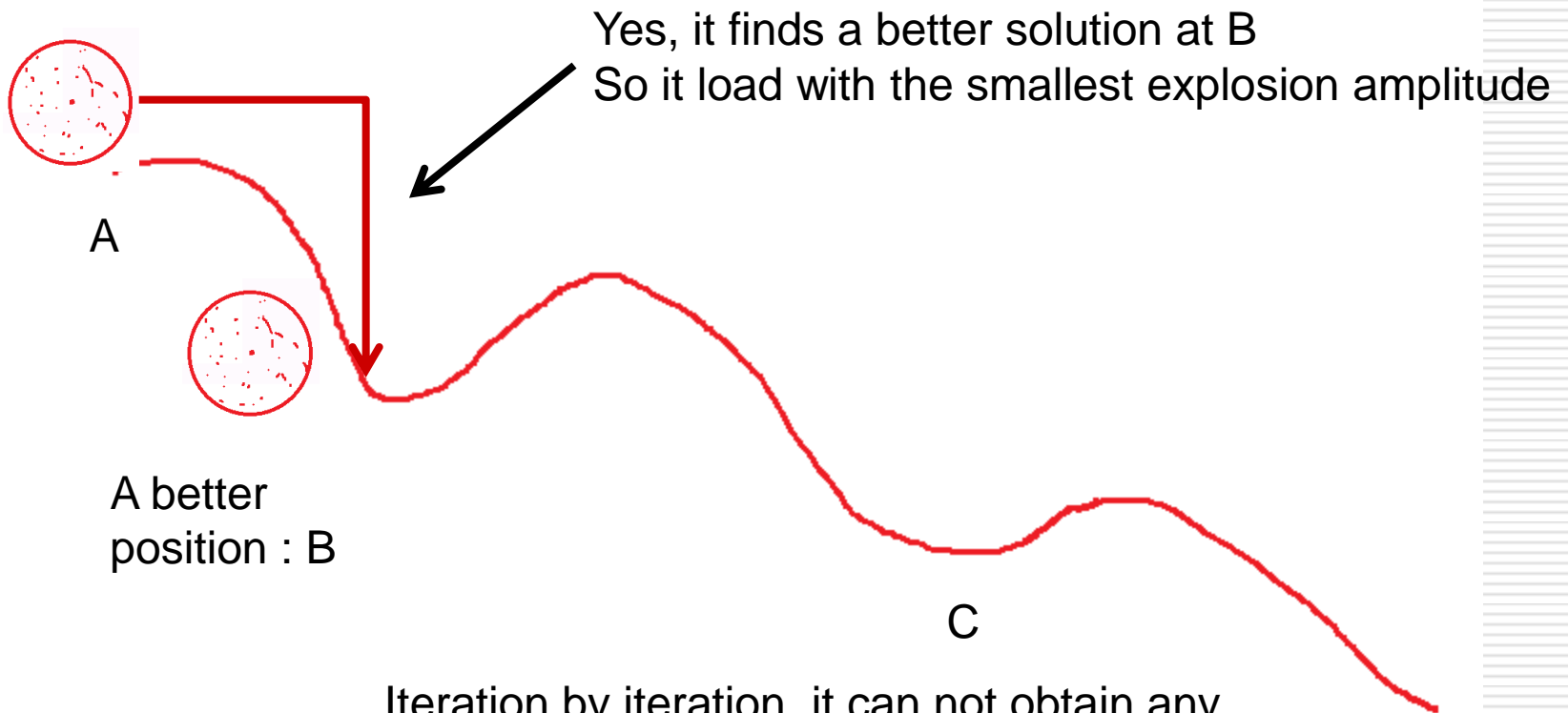
3.4 Two cases:



- ❑ the population of the swarm (excluding the ***core firework***) is too small.
- ❑ the search range of the core firework is too small

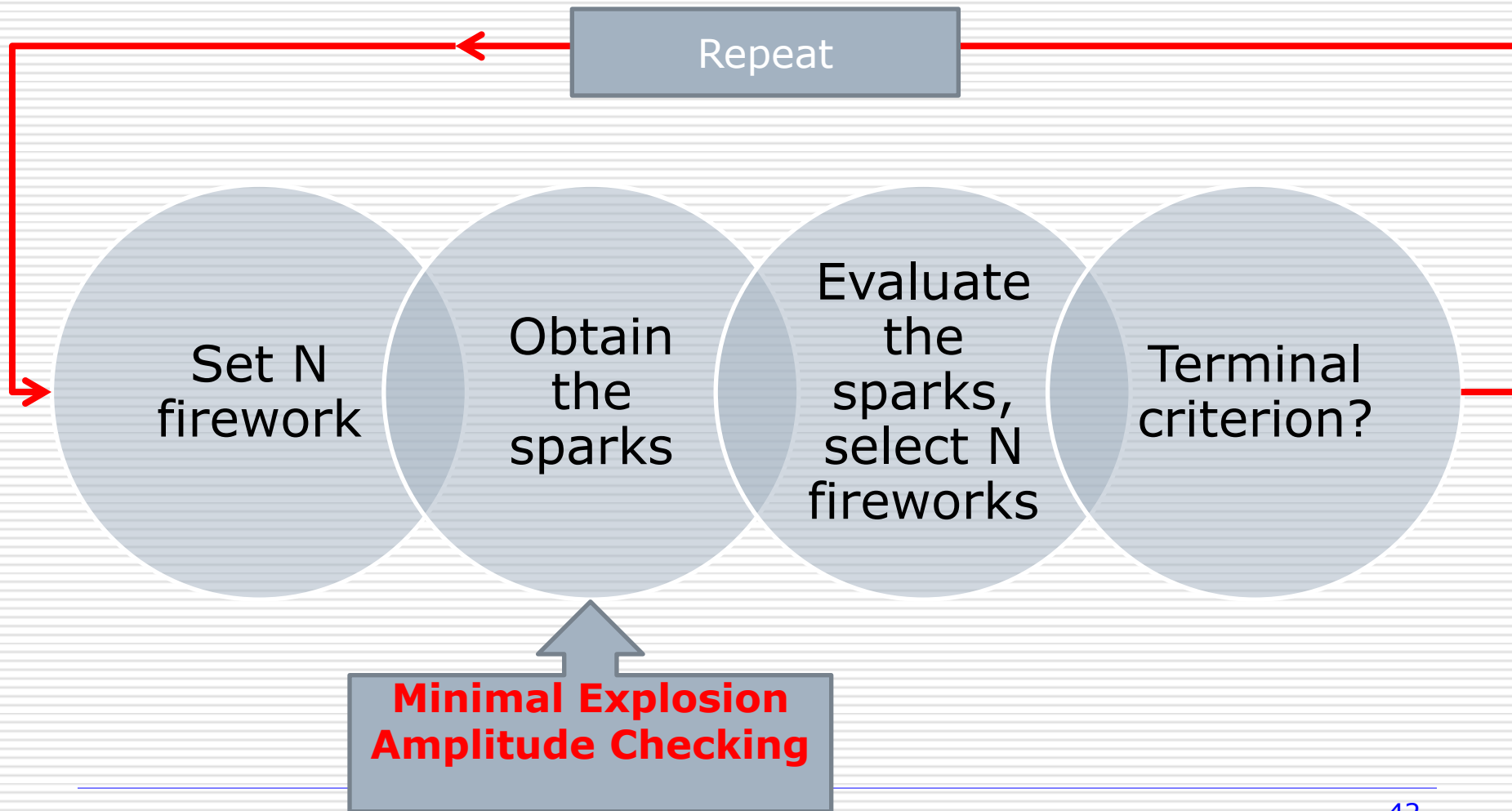
**LOW
DIVERSITY**

3.4 Dynamic minimal explosion amplitude strategy



Iteration by iteration, it can not obtain any improvement, so a bigger explosion amplitude Should be exploited till it reaches point C.

3.4 FA with dynamic minimal explosion amplitude



3.4 FA with dynamic explosion amplitude

□ 25 Benchmark Test Functions

TEST FUNCTIONS GROUP 2 - OPTIMA AND FEASIBLE RANGES
(UNI=UNIMODAL, MULTI=MULTIMODAL, SH=SHIFTED, RT=ROTATED, GB=GLOBAL ON BOUNDS, HC=HYBRID COMPOSITION, NM=NUMBER MATRIX)

No.	Description	Feasible Bounds	Optima Value
f_1	Sh Sphere	[-100, 100]	-450
f_2	Sh Schwefel 1.2		-450
f_3	Sh Rt Elliptic		-450
f_4	f_2 with Noise		-450
f_5	Schwefel 2.6 GB		-310
f_6	Sh Rosenbrock	[-100, 100]	390
f_7	Sh Rt Griewank	[0, 600]	-180
f_8	Sh Rt Ackley GB	[-32, 32]	-140
f_9	Sh Rastrigin	[-5, 5]	-330
f_{10}	Sh Rt Rastrigin	[-5, 5]	-330
f_{11}	Sh Rt Weierstrass	[-0.5, 0.5]	90
f_{12}	Schwefel 2.13	$[\pi, \pi]$	-460
f_{13}	Sh Expanded F8F2	[-3, 1]	-130
f_{14}	Sh Rt Scaffer F6	[-100, 100]	-300
f_{15}	HC Function	[-5, 5]	120
f_{16}	Rt HC Function 1		120
f_{17}	f_{16} with Noise		120
f_{18}	Rt HC Function 2		10
f_{19}	f_{18} with Basin		10
f_{20}	f_{18} with GB		10
f_{21}	Rt HC Function 3		360
f_{22}	f_{21} with NM		360
f_{23}	NC Rt f_{21}		360
f_{24}	Rt HC Function 4		260
f_{25}	f_{24} without Bounds		360

Here, f_1 - f_5 are simple uni-modal problems, f_6 - f_{14} are multi-modal problems with few local minima, f_{15} - f_{25} are highly complex hybrid-modal problems with a number of local minima.

3.4 FA with dynamic minimal explosion amplitude

□ Experimental Results of FA-DEA on f_1 - f_{14}

TABLE IV

RESULTS ON f_1 - f_{14} BENCHMARK TEST FUNCTIONS BY FA AND FA-DEA (M=MAX EVALUATION TIMES, $M_1 = 50000, M_2 = 250000, M_{3,4} = 1500000, M_5-M_{14}=150000$)

No.	PSO Mean \pm Std	FA Mean \pm Std	FA-DEA Mean \pm Std
f_1	-437.917 \pm 46.205	436.448 \pm 6.2110	-449.956 \pm 0.0000
f_2	-432.118 \pm 44.115	9506.19 \pm 2088.8	2168.23 \pm 1098.6
f_3	6.83e+6 \pm 7.29e+6	1116.96 \pm 724.07	-405.225 \pm 14.344
f_4	-160.973 \pm 158.42	7536.53 \pm 2186.1	5907.39 \pm 1889.5
f_5	5051.06 \pm 1980.4	938.752 \pm 1119.3	292.689 \pm 914.89
f_6	390.931 \pm 1.2703	802.257 \pm 259.84	894.834 \pm 837.99
f_7	-179.910 \pm 0.0443	11990.4 \pm 476.05	11788.1 \pm 531.59
f_8	-119.761 \pm 0.0611	16231.3 \pm 220.34	16228.0 \pm 287.17
f_9	-251.615 \pm 17.042	-265.640 \pm 19.353	-323.212 \pm 2.0897
f_{10}	-218.479 \pm 22.354	-271.118 \pm 13.113	-323.411 \pm 1.9395
f_{11}	-94.9889 \pm 0.9043	306.396 \pm 2.6608	303.040 \pm 1.3003
f_{12}	-434.730 \pm 48.938	377.667 \pm 15.064	-448.386 \pm 2.6561
f_{13}	-129.508 \pm 0.2479	25.0533 \pm 21.703	-40.3880 \pm 1.8349
f_{14}	-287.181 \pm 0.2455	-222.917 \pm 16.205	-287.802 \pm 2.9194

3.4 FA with dynamic minimal explosion amplitude

Experimental Results of FA-DEA on f_{15} - f_{25}

TABLE V

RESULTS ON f_{15} - f_{25} BENCHMARK TEST FUNCTIONS BY FA AND FA-DEA (M=MAX EVALUATION TIMES, $M_{15} - M_{25} = 150000$)

No.	PSO Mean \pm Std Percentage	FA Mean \pm Std Percentage	FA-DEA Mean \pm Std Percentage
f_{15}	310.560 \pm 100.24 0/20	358.900 \pm 200.79 0/20	257.015 \pm 202.59 11/20
f_{16}	256.132 \pm 16.535 0/20	345.319 \pm 186.49 0/20	295.922 \pm 217.05 11/20
f_{17}	267.267 \pm 17.062 0/20	346.059 \pm 180.96 0/20	312.323 \pm 216.04 10/20
f_{18}	781.612 \pm 118.15 0/20	249.643 \pm 190.14 0/20	188.298 \pm 214.31 7/20
f_{19}	780.762 \pm 104.16 0/20	298.746 \pm 193.96 0/20	173.548 \pm 207.61 8/20
f_{20}	765.982 \pm 86.237 0/20	193.664 \pm 172.44 0/20	182.559 \pm 208.76 9/20
f_{21}	949.150 \pm 193.30 0/20	598.317 \pm 207.12 0/20	570.952 \pm 213.54 6/20
f_{22}	1140.31 \pm 25.381 0/20	505.953 \pm 139.23 0/20	545.139 \pm 215.46 9/20
f_{23}	1085.95 \pm 229.70 0/20	546.338 \pm 156.31 0/20	500.896 \pm 198.10 10/20
f_{24}	1172.81 \pm 90.035 0/20	490.309 \pm 194.91 0/20	363.150 \pm 152.02 9/20
f_{25}	641.631 \pm 3.6937 0/20	512.075 \pm 209.82 0/20	406.561 \pm 211.35 10/20

Performance Comparison on Various Dimensions

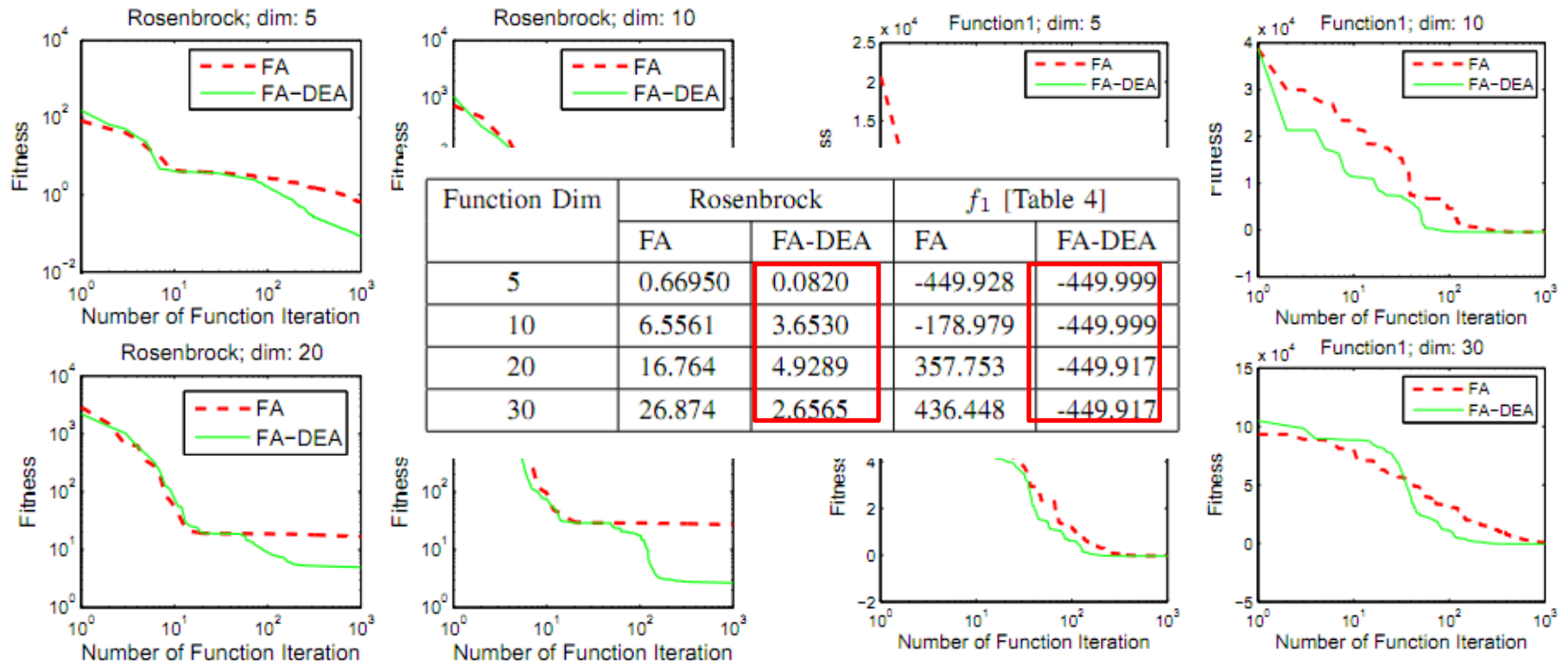


Fig. 10. Experiments on Rosenbrock and $Dim = 5, 10, 20$ and 30

Fig. 11. Experiments on f_1 and $Dim = 5, 10, 20$ and 30

4. Application Research

- 4.1 FA for NMF
- 4.2 FA for Clustering

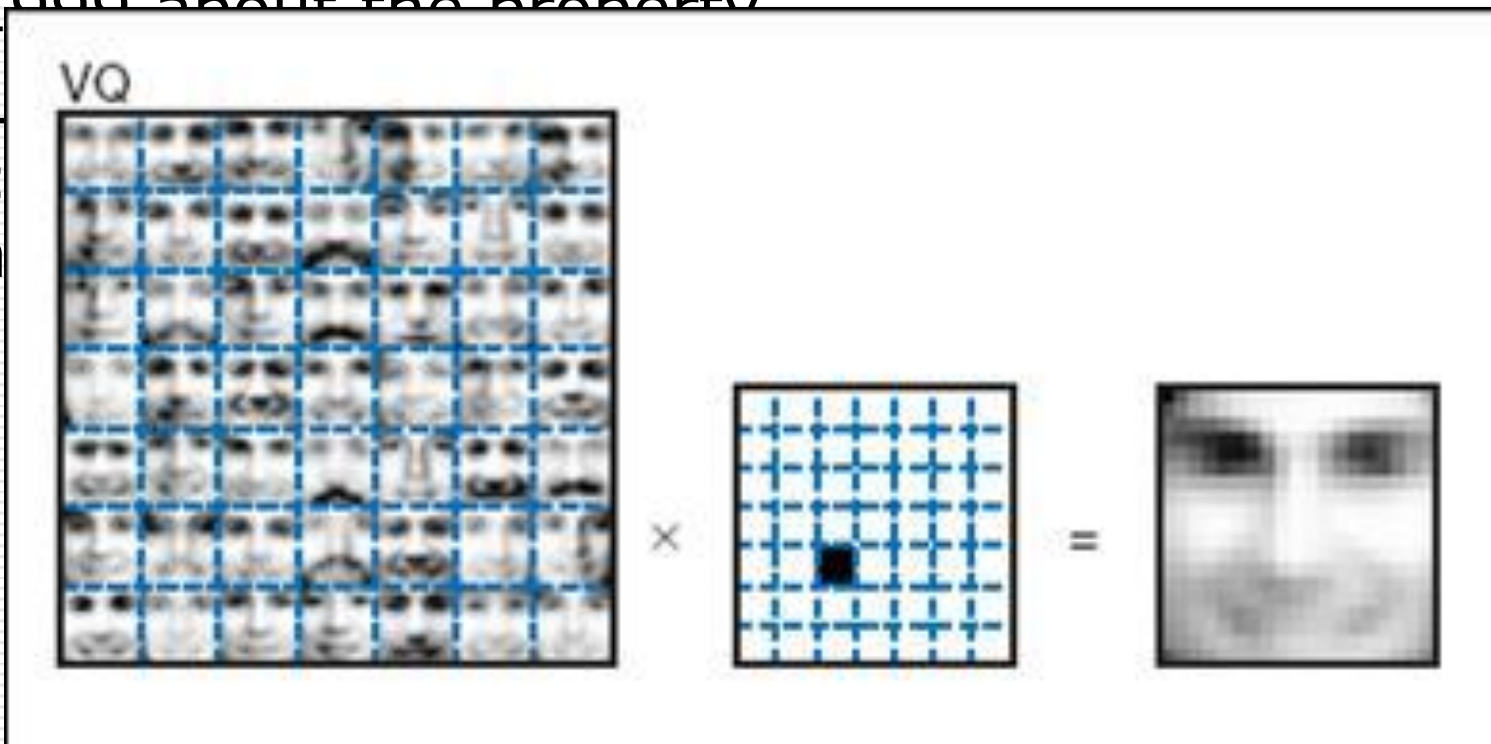
4.1 FA for NMF

- 4.1.1 NMF description
- 4.1.2 Algorithm for NMF Computing
- 4.1.3 Experimental results

4.1.1 NMF description

- Lee and Seung publish a paper on Nature in 1999 about the property

- L
C
a



4.1.1 NMF description

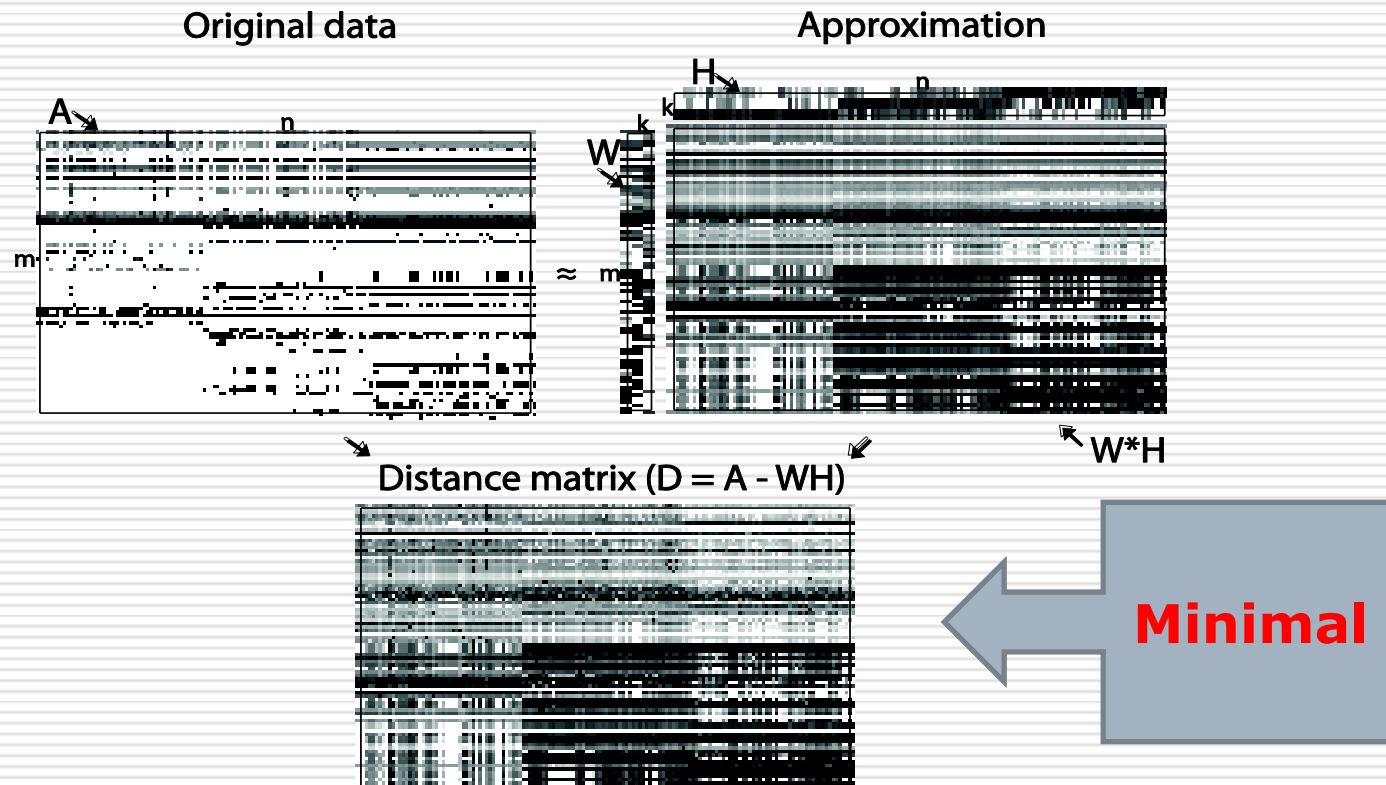


Figure - Scheme of very coarse NMF approximation with very low rank k . Although k is significantly smaller than m and n , the typical structure of the original data matrix can be retained (note the three different groups of data objects in the left, middle, and right part of A).

4.1.1 NMF description

- Mathematically, we consider the problem of finding a “good” (ideally the global) solution of an optimization problem with bound constraints .

4.1.1 NMF description

- The nonlinear optimization problem underlying NMF can generally be stated as

$$\min_{W, H} f(W, H) = \min_{W, H} \frac{1}{2} \|A - WH\|_F^2 .$$

4.1.2 Algorithm for NMF Computing

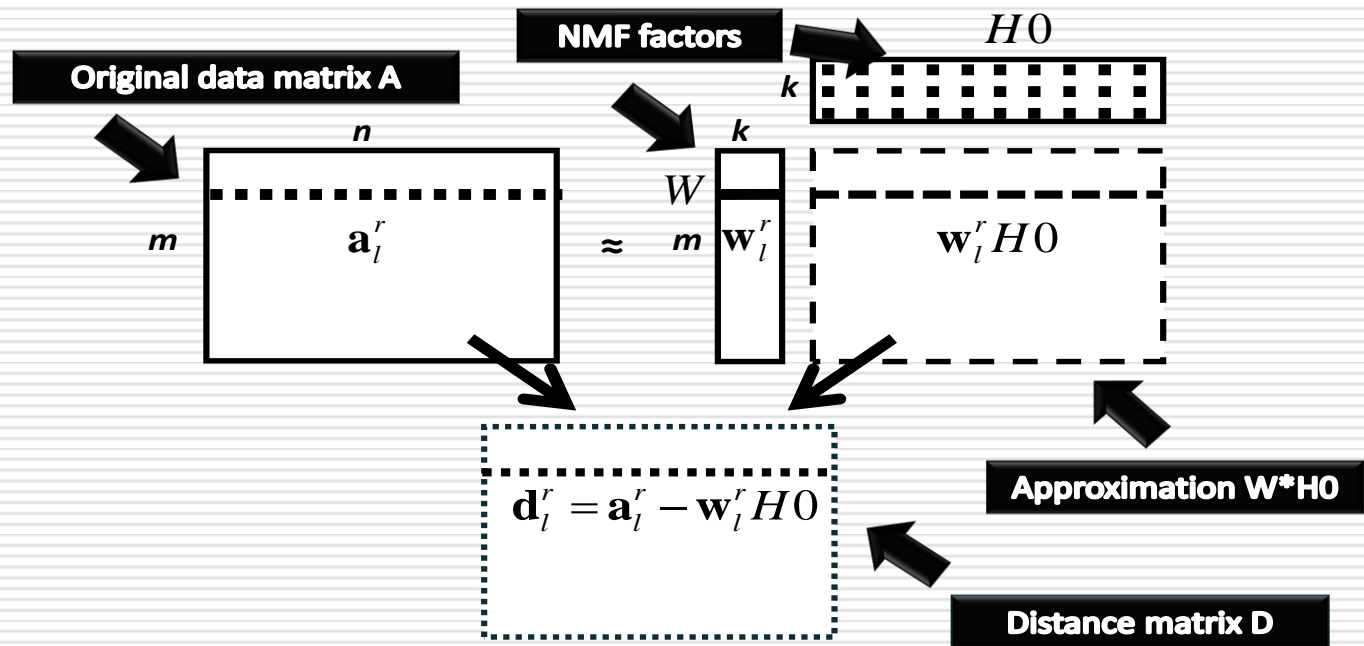


Figure – Illustration of the optimization process for row l of the NMF factor W . The l^{th} row of A (\mathbf{a}_l^r) and all columns of $H0$ are the **input** for the optimization algorithms. The **output** is a row-vector \mathbf{w}_l^r (the l^{th} row of W) which minimizes the norm of \mathbf{d}_l^r , the l^{th} row of the distance matrix D . The norm of \mathbf{d}_l^r is the fitness function for the optimization algorithms (minimization problem).

4.1.2 Algorithm for NMF Computing

□ *General structure of NMF algorithms*

-
- 1: given matrix $A \in \mathbb{R}^{m \times n}$ and $k \ll \min\{m, n\}$:
 - 2: **for** $rep = 1$ to $maxrepetition$ **do**
 - 3: $W = \text{rand}(m, k)$;
 - 4: $(H = \text{rand}(k, n))$;
 - 5: **for** $i = 1$ to $maxiter$ **do**
 - 6: perform algorithm specific NMF update steps
 - 7: check termination criterion
 - 8: **end for**
 - 9: **end for**
-

4.1.2 Algorithm for NMF Computing

□ NMF Initialization

```
1: Given matrix  $A \in \mathbb{R}^{m \times n}$  and  $k \ll \min\{m, n\}$ ;  
2:  $H0 = \text{rand}(k, n)$ ;  
3: % Compute in parallel  
4: for  $i = 1$  to  $m$  do  
5:   Use SIO to find  $\mathbf{w}_i^r$  that minimizes  $\|\mathbf{a}_i^r - \mathbf{w}_i^r H0\|_F$ , (min  $\|\cdot\|_F$  of row  $i$  of  $D$ );  
6: end for;  
7: % Gather  
8:  $W = [\mathbf{w}_1^r; \dots; \mathbf{w}_m^r]$ ;  
9: % Compute in parallel  
10: for  $j = 1$  to  $n$  do  
11:   Use SIO to find  $\mathbf{h}_j^c$  that minimizes  $\|\mathbf{a}_j^c - W\mathbf{h}_j^c\|_F$ , (min  $\|\cdot\|_F$  of col  $j$  of  $D$ );  
12: end for  
13: % Gather  
14:  $H = [\mathbf{h}_1^c, \dots, \mathbf{h}_n^c]$ ;
```

4.1.2 Algorithm

Iterative Optimization

```

1: for  $iter = 1$  to  $maxiter$  do
2:   % perform MU specific update steps
3:    $W = W \cdot (AH^\top) ./ (WHH^\top + \varepsilon);$ 
4:    $H = H \cdot (W^\top A) ./ (W^\top WH + \varepsilon);$ 
5:   if ( $iter < m$ ) then
6:     % Update rows of  $W$  ++++++
7:      $\mathbf{d}_i^r$  is the  $i^{th}$  row vector of  $D = A - WH$ ;
8:      $[Val, IX\_W] = sort(norm(\mathbf{d}_i^r), 'descend');$ 
9:      $IX\_W = IX\_W(1:c);$ 
10:    % Compute in parallel
11:     $\forall i \in IX\_W:$ 
12:      Use SIO to find  $\mathbf{w}_i^r$  that minimizes  $\|\mathbf{a}_i^r - \mathbf{w}_i^r H\|_F$ ;
13:    % Gather
14:     $W = [\mathbf{w}_1^r; \dots; \mathbf{w}_m^r];$ 
15:    % Update columns of  $H$  ++++++
16:     $\mathbf{d}_j^c$  is the  $j^{th}$  column vector of  $D = A - WH$ ;
17:     $[Val, IX\_H] = sort(norm(\mathbf{d}_j^c), 'descend');$ 
18:     $IX\_H = IX\_H(1:c);$ 
19:    % Compute in parallel
20:     $\forall j \in IX\_H:$ 
21:      Use SIO to find  $\mathbf{h}_j^c$  that minimizes  $\|\mathbf{a}_j^c - W\mathbf{h}_j^c\|_F$ ;
22:    % Gather
23:     $H = [\mathbf{h}_1^c, \dots, \mathbf{h}_n^c];$ 
24:     $c = c - \Delta c;$ 
25:   end if
26: end for

```

4.1.2 Datasets for NMF Computing by FA

- ❑ *DS-RAND* is a randomly created, fully dense matrix which is used in order to provide unbiased results. To evaluate the proposed methods in a classification context we further used two data sets from the area of email classification (spam/phishing detection).
- ❑ Data set *DS-SPAM1* consists of 3000 e-mail messages described by 133 features, divided into three groups: spam, phishing and legitimate email.
- ❑ Data set *DS-SPAM2* is the *spambase* data set taken from (Kjellerstrand 2011) which consists of 1813 spam and 2788 non-spam messages. *DS-SPAM1* represents a ternary classification problem; *DS-SPAM2* represents a typical binary classification problem.

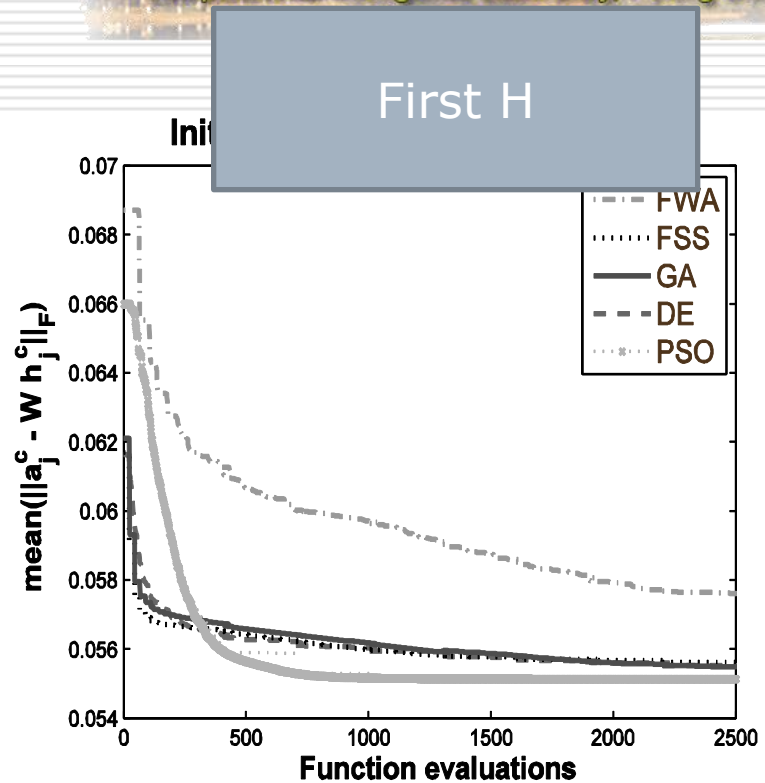
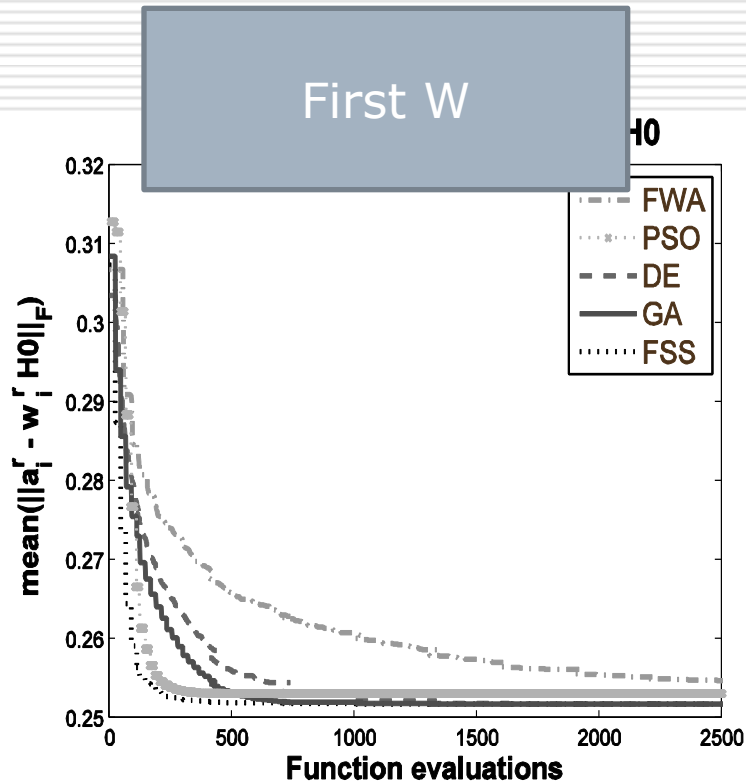


Figure 4.1.1 – Left hand-side: average approximation error per row (after initializing rows of W). Right hand-side: average approximation error per column (after initializing of H). NMF rank $k = 5$. Legends are ordered according to approximation error (top = worst, bottom = best).

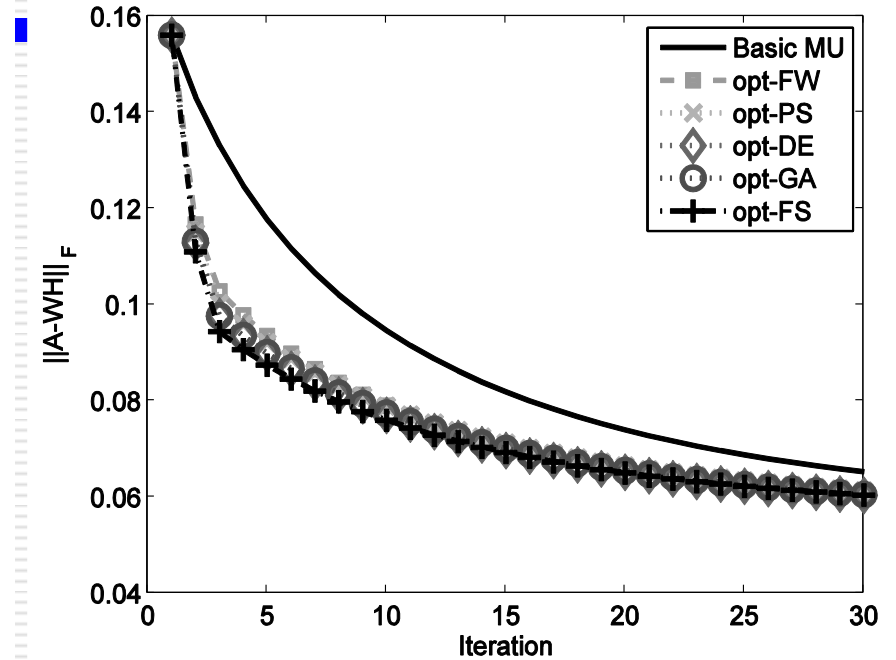
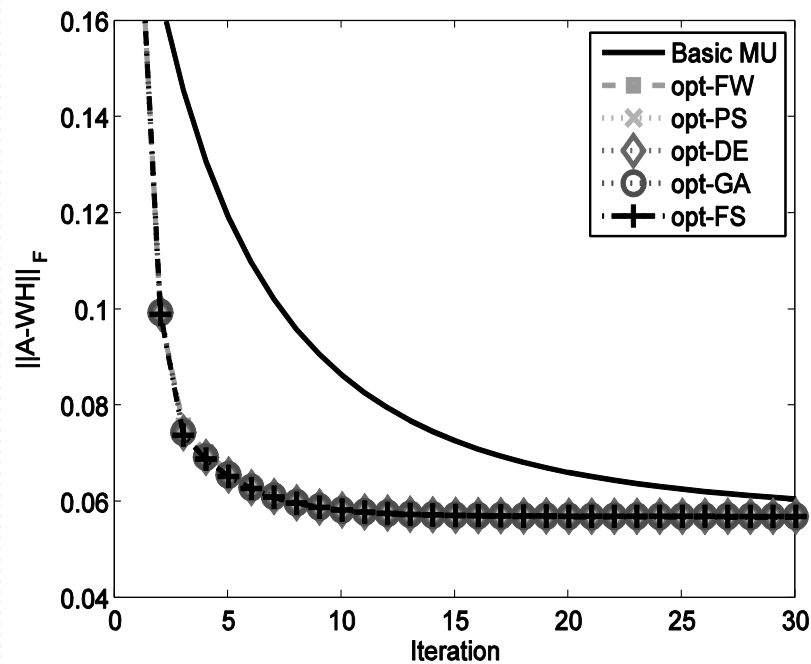


Figure 4.1.2 Accuracy per Iteration when updating only the row of W , $m=2$, $c=20$. Left: $k=2$, right: $k=5$

Experiments results of NMF

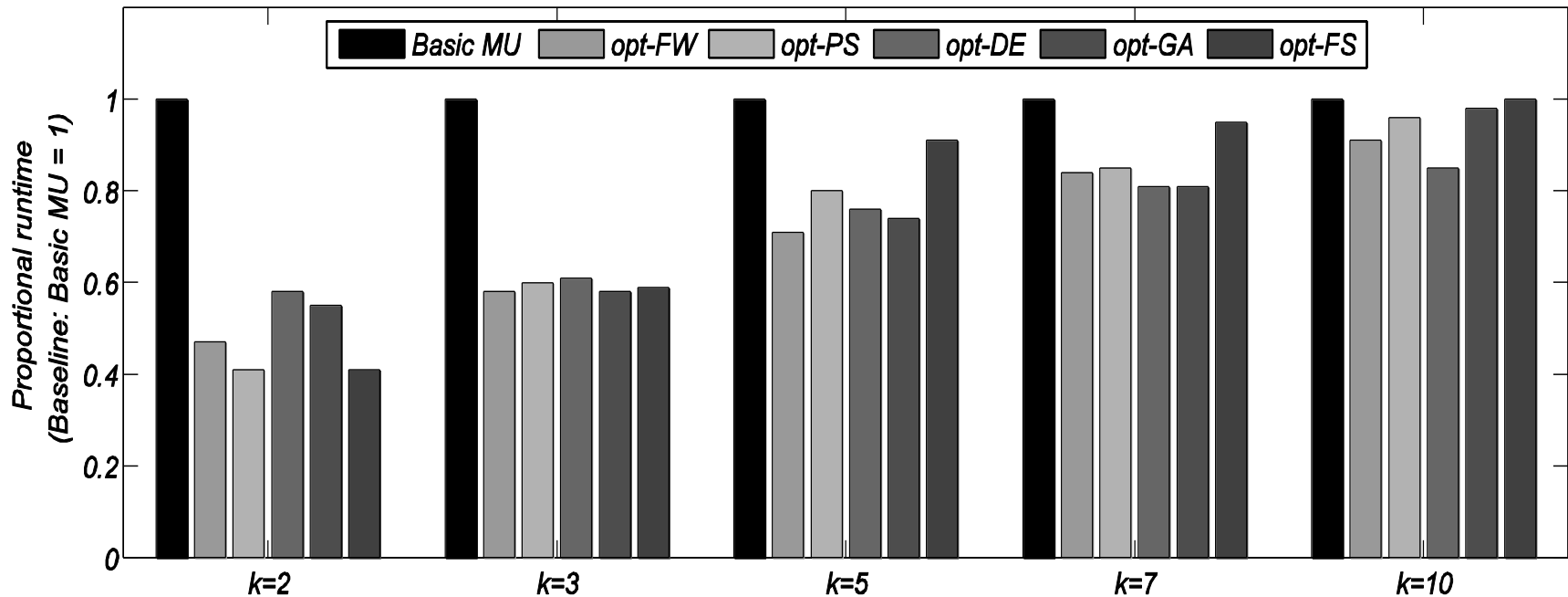


Figure 4.1.3 – *Proportional runtimes for achieving the same accuracy as basic MU after 30 iterations for different values of k when updating only the rows of W . ($m=2$, $c=20$)*

4.2 FA for Document Clustering

- 4.2.1 Document Clustering Description
- 4.2.2 Dataset
- 4.2.3 Experimental Result

4.2.1 Document Clustering description

- ❑ Automatically group the related documents into clusters.
 - Example
 - Medical documents
 - Legal documents
 - Financial documents
- ❑ If a collection is well clustered, it is much more efficient to search only the cluster that will contain relevant documents .

4.2.2 Dataset: Newsgroups-18828

predecessor: Newsgroups

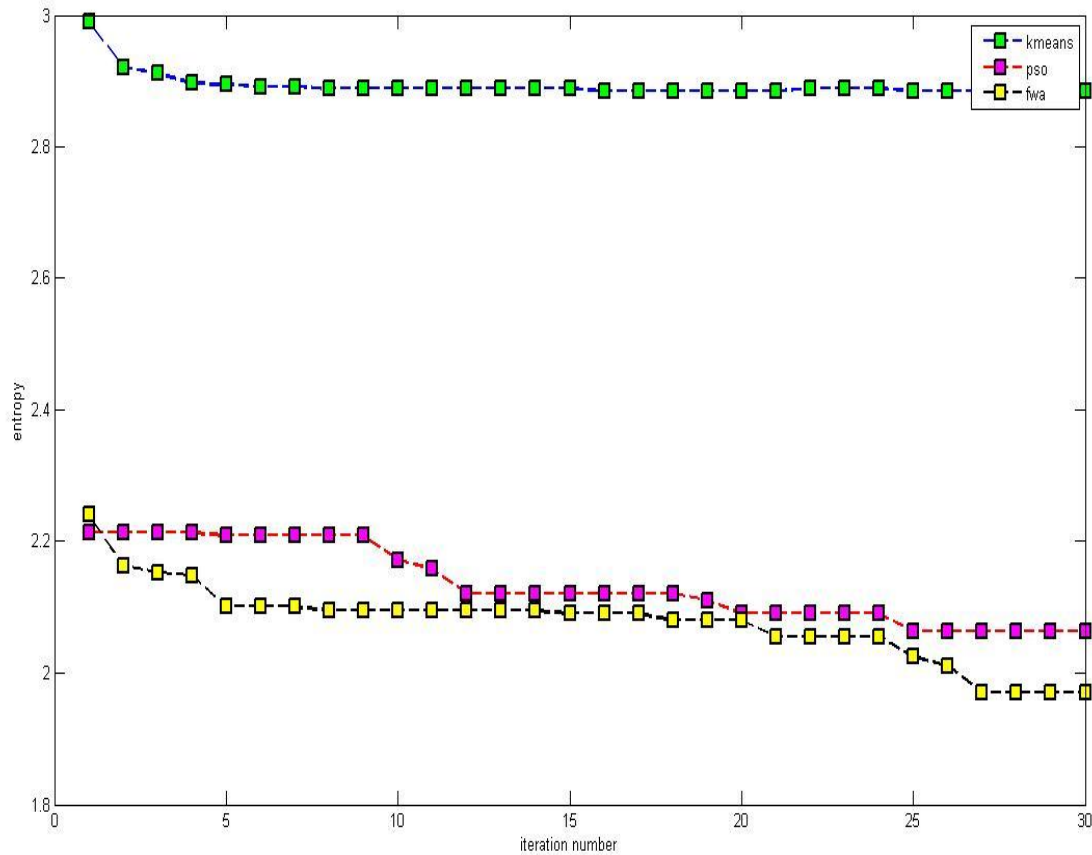
- ❑ The 4.5% document belongs to the two or more than two news group
- ❑ The remaining documents only belongs to one newsgroup

modification : Jason Rennie from MIT do some necessary processings to Newsgroups, so that each document belongs to only one News group

characters: A total of 18828 documents, all documents belong to 20 different new newsgroups

- widely used in document classification and clustering

4.2.3. Experimental Result



Concluding Remarks

- By mimicking the explosion process of fireworks in our festivals, the so-called **FA** is proposed and implemented for function optimization, with a promising performance against **Clonal PSO** and **Standard PSO**.
- **FA-DEA**, by introducing an explosion amplitude control strategy, has shown its great advantages in hybrid-modal problems, which can achieve the optima 50% on average.
- The use of Firework Algorithm to solve practical engineering and scientific problems like clustering and NMF has gained great success compared to traditional methods.

Acknowledgment

- This work was supported by National Natural Science Foundation of China (NSFC), under Grant No. 60875080, 60673020, and partially supported by the 863 Program with Grant No. 2007AA01Z453.
- Greatly appreciate my Ph.D. students, ***Zhu Yuanchun*** and ***Zheng Shaoqiu***, for their excellent contributions and extensive experiments.
- Many thanks go to my Postdoc ***Dr. Andreas Janecek*** from Univ. of Vienna, and my M.S. students, ***Yang Xiang*** and ***Chao Rui*** for their significant application researches on FA.

Reference

- [1] Y. Tan, Y. Zhu (2010). Fireworks algorithm for optimization. **ICSI 2010**, Part I, Springer LNCS 6145, pp. 355-364
- [2] S.Q. Zheng, Y. Tan (2011). Fireworks Algorithm with Dynamic Explosion Amplitude. **Submitted to IEEE CEC 2012.**
- [3] A. Janecek and Y. Tan, "Using Population Based Algorithms for Initializing Nonnegative Matrix Factorization," **Proc. of The Second International Conference on Swarm Intelligence (ICSI'2011)**, June 12-15, Chongqing, China, 2011. Part II, LNCS 6729, Springer-Verlag, 2011, pp.307-316.
- [4] A. Janecek and Y. Tan, "Iterative Improvement of the Multiplicative Update NMF Algorithm using Nature-inspired Optimization," **The 7th International Conference on Natural Computation (ICNC'11)**, 26-28 July 2011, Shanghai, China, Springer LNCS, pp.1668 - 1672
- [5] A. Janecek, Y. Tan, " Swarm Intelligence for Non-negative Matrix Factorization," **International Journal of Swarm Intelligence Research (IJSIR)**, November 2011 (In press)
- [6] X. Yang, Using Fireworks Algorithm for Document Clustering, **Bachelor Thesis, Peking University**, 2011.

Thank you

***Any
Question ?***

